THE UNIVERSITY of EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

The Impact of Dynamic Soundfield on Delivering Improvements in Educational Attainment and Closing the Attainment Gap with Young Learners in Mainstream Primary School

Brian B Shannan

Submitted for the Degree of Doctor of Philosophy - The University of Edinburgh - 2019

Abstract

Introduction

Young people in classrooms are generally listening to speech in an acoustically complex environment composed of long reverberation times and multiple competing voices generated by the room occupants. The challenges of listening to conversation when there are several competing voices is known as the ‘cocktail party’ or ‘multi-talker listening’ effect. Many of the cues that help to filter speech in noise are sometimes inaccessible to young people due to an auditory system that is going through a process of maturation; young people have limited life experience and an underdeveloped knowledge of linguistic rules. Education systems throughout the UK and beyond have gone through a period of reform, with the traditional didactic teaching methods supplemented by more interactive and student-focused learning which are commonly associated with higher noise levels. Several studies have shown that noise levels have a detrimental effect on learning and cognitive function. At the same time, concerns over the general attainment levels in schools and the gap between learners from the most and least deprived areas have come to the forefront of government policy.

Aims

The primary purpose of this research was to evaluate the effects of dynamic soundfield technology on educational attainment with young primary three learners and to determine if its efficacy is moderated by the speech weighted C_{50} properties of the classrooms. Dynamic soundfield is a technology that monitors the level of background noise in a classroom and adapts the level of gain provided to the teachers’ voice to enhance speech intelligibility in noise. Another primary aim was to explore and evaluate the effects of dynamic soundfield on learners from Scottish Index of Multiple Deprivation (SIMD) 1 and 5 to discover if it contributed to closing the attainment gap between learners from the most and least deprived areas of Scotland. One of the secondary aims of this study was to measure the acoustic properties of the occupied and unoccupied research classrooms and to establish quantitative data on the overall use of the dynamic soundfield systems in each intervention classroom. In addition, the young learners’ and teachers’ experience of noise when exposed to the intervention was compared to the control. The findings from the research are discussed with reference to the literature and in relation to implications for practice.

Methods

This research was a longitudinal repeat measure study design. The 495 primary three learners from 13 schools were allocated to control and intervention classrooms through concealed randomisation. Overall, there were 25 research classrooms, 13 intervention classrooms were fitted with dynamic soundfield (278 learners) and 12 acted as a control (217 learners). All the participants completed the Achievement

for Excellence (Interactive Computerised Assessment System) (AfE (InCAS)) suite of adaptive assessments at the start and end of the study. These were administered and marked by the CEM at the University of Durham who were blind to the intervention.

Results

Comparisons of the pre-test and post-test Developed Ability and Mental Arithmetic module scores revealed that dynamic soundfield systems were primarily effective in classrooms with good C_{50} values. In the data handling subtest, the dynamic soundfield intervention produced significant improvements in classrooms with excellent acoustics for speech. In reading, the intervention classrooms categorised as fair and good for speech clarity demonstrated significantly higher learner scores in word recognition, comprehension and reading compared to the control. Only the good C_{50} intervention classrooms demonstrated a significant improvement in word decoding. Results from a series of two-way mixed ANOVAs showed that learners from the most deprived quintile exposed to dynamic soundfield in classrooms with good C_{50} values gained a significant benefit in the Developed Ability module and subtests. Learners from the least deprived quintile, in classrooms with excellent C_{50} values, showed a statistically significant improvement in the General Mathematics subtests of Data Handling and Measure, Shape and Space. Learners from both SIMD 1 and 5 quintiles demonstrated a significant benefit in the Reading module and Word Recognition subtests. Only SIMD 1 learners showed a significant advantage of dynamic soundfield in the word decoding subtest, which was primarily effective in good C_{50} classrooms. In the Non-Verbal Ability and Word Recognition subtests, there was a significant reduction in the gap between SIMD 1 learners in the intervention classroom and SIMD 5 learners in the control.

The views and experiences of learners towards classroom noise and the use of dynamic soundfield were explored quantitively: noise questionnaires were completed by all participants. Logistic regression model analysis revealed that learners exposed to dynamic soundfield could hear the class teacher more easily compared to the control. Overall, the results indicate that dynamic soundfield may be regarded as a cost-effective intervention strategy that could contribute to raising attainment and reducing inequality in Scottish education.

Lay Abstract

Introduction

Young people find listening to speech in noise more challenging than adults. In a classroom each person talking creates a sound, in addition, there are multiple reflections of the original sound and these can overlap and hide speech. Young people when listening in a classroom require to separate the original sound from the reflected sounds to hear speech effectively. This can be challenging for young learners as they do not have the same life experience or knowledge of language compared to an adult. Also, the auditory system continues to develop into the teenage years. Changes in educational practice have resulted in an increase in learning activities known to produce noise. Research has shown that noise can have a negative effect on learning. At the same time, the government is also committed to raising the level of achievement in schools and cutting the gap between young people living in the most and least deprived areas of the country.

Aims

This research aims to find out if dynamic soundfield helps to improve learning outcomes amongst primary three learners. Dynamic soundfield is a system that amplifies the voice of the teacher making it easier to hear in noise. The system checks the noise levels in the classroom and increases the level of the teachers' voice as noise levels get higher. The research also aims to find out if learners from the most and least deprived areas benefit equally from the system. The research will also explore whether learners from the most deprived areas who had dynamic soundfield fitted in the classroom did better than the learners from the least deprived areas who had no access to a soundfield. Measurements of the classrooms were made to see how good they were for hearing speech. Also, we checked to see how often the dynamic soundfield was used during the research. We also asked the teachers and learners about their experience of noise in the classroom. We also asked how well the dynamic soundfield system worked.

Methods

This research lasted for a school year. There were 495 primary three learners from 13 schools who took part in the study. There were 25 research classrooms. In 13 classrooms we fitted the dynamic soundfield system. There were 278 learners in these classrooms. In 12 classrooms we fitted no equipment. There were 217 learners in these classrooms. All the learners completed a set of tests on a computer. These were completed before we started our study and when we finished our study. These tests were marked by the University of Durham who did not know which learners had the dynamic soundfield in their classroom.

Results

The classrooms that were good for hearing speech and were fitted with a dynamic soundfield had better results than the classrooms that were good for speech but did not have a soundfield. This was the case for all the Developed Ability tests. It was also the case for Mental Arithmetic. The classrooms that were excellent for hearing speech and had a dynamic soundfield fitted performed better in the data handling test than the classrooms that were excellent for speech but had no soundfield. In reading, word recognition, and listening comprehension it was the classrooms that were fair and good for hearing speech that had a dynamic soundfield that performed the best. Learners in classrooms good for speech, fitted with a dynamic soundfield did the best in word decoding. Learners from the most deprived areas in classrooms fitted with dynamic soundfield that were good for hearing speech performed very well in the Developed Ability and word decoding test. Learners from the least deprived areas in classrooms with excellent speech showed a significant improvement in the Data Handling and Measure, Shape and Space tests. Both groups gained a significant benefit in the Reading module and Word Recognition tests. In the Non-Verbal Ability and Word Recognition tests, there was a significant reduction in the gap between learners from the most deprived areas with dynamic soundfield fitted in their classroom and the learners from the least deprived that had no soundfield fitted.

Noise questionnaires were completed by all learners. The results showed that learners in classrooms fitted with dynamic soundfield could hear the class teacher more easily compared to the classrooms with no soundfield. Overall, the results show that dynamic soundfield may be effective at improving learning in schools and may help reduce that gap between learners from the most and least deprived areas.

Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where stated otherwise by reference, or acknowledgement, the work presented is entirely my own.

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated. This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author. The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.

[redacted]

1st September 2019

Acknowledgements

I am indebted to all the young people and school staff who participated in this research. Without their time, flexibility and patience this study would not have been possible.

I am extremely grateful to all the staff of Fife Council who offered advice, support, and guidance. My deepest thanks are also extended to all the staff of the Sensory Support (Deaf) Service whose contribution was invaluable.

It has been a pleasure to have been supervised by Professor John Ravenscroft and Professor Murray Campbell. The support and guidance they provided through each stage of this study were invaluable. Thank you both for your time, knowledge, patience and humour. At different times during this study, they provided me with support which included at least one of these attributes. I am sincerely grateful.

Abbreviations

AfE (InCAS)	Achievement for Excellence (Interactive Computerised Assessment System)
ASN	Additional Support Need
CEM	Centre for Evaluating and Monitoring
CfE	Curriculum for Excellence
C_{50}	Speech clarity
ENR	Equivalent Noise Reduction
IDL	Interdisciplinary Learning
ILD	Interaural level difference
ITD	Interaural time difference
LIFE	Listening Inventory for Education
RT	Reverberation time
SIMD	Scottish Index of Multiple Deprivation
SNR	Signal to Noise Ratio
STI	Speech Transmission Index
T_{mf}	Mid-frequency reverberation time

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Defining the area of inquiry

1.1.1 Noise and reverberation

Boothroyd (2004) makes two assumptions about school classrooms. Firstly, a classroom is an environment where learning occurs and secondly, the majority of learning is mediated through spoken language. Listening to spoken language in a classroom generally occurs in an imperfect environment, one in which there are multiple competing voices, long mid-frequency reverberation times (T_{mf}) and interference from environmental noise sources, both external and internal. External environmental noise can originate from a range of sources including roads, trains, aircraft, and industry depending on the location of schools. Building services such as boilers and fans contribute to the internal noise levels and the combined acoustical impact of both these external and internal sources defines the unoccupied ambient sound environment (Institute of Acoustics, 2015). The recommended unoccupied ambient noise levels for classrooms in England and Wales is 35 L_{Aeq} (Department of Education and Skills, 1993). A number of studies, both nationally and internationally have demonstrated that the acoustic properties of school classrooms and ambient noise levels do not comply with the recommended guidelines (Hodgson, 1994, Knecht et al., 2002, Yee Choi and McPherson, 2005, Shield and Dockrell, 2008).

In general, external noise only affects internal noise levels in occupied classrooms when the room is quiet (Shield and Dockrell, 2008). The noise levels in an occupied classroom are primarily influenced by the number of learners in the room, the task that the learners are involved in, age of the learners, classroom management and the acoustic properties of the room (Klatte et al., 2010a, Shield and Dockrell, 2004, Shield and Dockrell, 2008). Shield and Dockrell (2004) studied 140 classrooms in 16 schools in London and found that age and learning activity had a significant impact on overall noise levels. The classroom noise levels were in excess of 70 L_{Aeq} in the nursery until Year 1 (age 5-6) when this reduced slightly before steadily increasing in Years 5 (age 9-10) and 6 (10-11). Most sound levels are not steady and fluctuate over time. L_{Aeq} is the most widely used unit to describe sound levels that vary over time and is an energy average (over a defined period of time) expressed as a single A-weighted decibel value (Institute of Acoustics, 2015). The research also found an association between the level of internal noise and the activity in which the learner was engaged. Six different learning activities were identified in the classrooms and these ranged from silent reading to learners working in groups. The difference in noise levels between the two activities differed by approximately 20 dB L_{Aeq} with silent reading/class test being recorded at 56.3 L_{Aeq} and group work with learners moving around the room at 76.8 L_{Aeq}.

Reverberation also acoustically impedes speech intelligibility in a classroom by introducing additional noise and temporal and spectral distortion to the speech signal of interest (Berg et al., 1996). Reverberation is the persistence of multiple and repeated reflections from the surface of the room and the RT is the time in seconds required for sound pressure at a specific frequency to decay after the sound source has stopped (Klatte et al., 2010b). RT_{60} is the commonly used abbreviation for the time it takes for the sound pressure level to decrease by 60dB. RT_{60} is strongly influenced by the volume, shape, construction materials and finishes within a room. A ‘dry’ environment is one lacking in acoustic reflections and is objectively characterised by a very short RT_{60}. A ‘live’ room will generally have a higher amount of hard surfaces from which sound is readily reflected and is objectively characterised by long RT_{60} (Institute of Acoustics, 2015).

Speech signals heard by the listener are a mixture of direct signal and time-delayed reflections. For learners with normal hearing, if the reflected sound arrives within 50 milliseconds of the direct sound it can enhance speech intelligibility (Bradley et al., 2003). However, reflections arriving more than 75 milliseconds after the direct sound result in a smearing of the speech signal which masks the speech sound and reduces the clarity of the speech and thereby intelligibility (Boothroyd, 2004).

1.1.2 The cocktail party problem

.

The seminal work of Cherry (1953) addressed the complex and multidimensional problem of speech recognition in noise with a simple question: ‘How do we recognise what one person is saying when others are speaking at the same time?’ This is commonly known as the Cocktail Party or multi-talker problem. Noise from environmental sources and concurrent speech in a room can interfere with the speech recognition process as it masks speech sounds which result in the learner receiving either partial or inaudible auditory information. In general terms, masking refers to a degrading of the target speech signal by other sound sources (Durlach, 2006). Masking can occur both in the peripheral auditory system and the central auditory mechanism (Cooke et al., 2008).

To achieve a release from masking, the learner requires to extract the target signal of interest from the competing and overlapping noise. Garadat and Litovsky (2007) observe that to achieve a high level of performance the learner in the classroom must segregate the multiple signals into their component parts to make sense of the ‘what’ and ‘where’ of their auditory environment. This involves processes such as auditory attention, binaural hearing and auditory scene analysis (Bronkhorst, 2000).

1.1.3 Issue 1 – the effect of noise and reverberation on attainment

Several studies have demonstrated the negative effect that noise has on educational attainment. For example, research in 16 London schools compared noise levels against Standard Assessment Test results at Key Stage 1 (7 years old) and 2 (11 years old) and found that external and internal noise present in a classroom affected performance. In general, the correlation between external noise and test scores were stronger for the older learners, suggesting that intermittent individual noise events such as motorbikes or lorries passing has a greater impact on this age group. The internal background noise levels in occupied classrooms impacted on all test results apart from Key Stage 1 Spelling and Key Stage 2 Mathematics. Key Stage 2 English was most effected by noise at both L_{Aeq} and L_{A90} levels (Shield and Dockrell, 2008).

This supports previous studies that examined the effect of noise on reading skills. A cohort model was used to assess the pre-reading skills of 90 pre-school young people (M=4 years and 7 months) over a two-year period. The participants were tested on three cognitive pre-reading measures: number and letter recognition, letter-sound correspondence and rhyming. In year one testing was performed in noisy conditions and in the second-year sound-absorbent panels in the ceiling were fitted. In the treated rooms the students performed better on recognition of numbers, letters and simple words but there was no significant difference in letter-sound correspondence and rhyming words. In addition the students in the noisy classrooms took longer to solve a puzzle immediately after working on an unsolvable one (Maxwell and Evans, 2000).

Long RT is a further factor that impedes attainment. In a field study involving 398 students in 17 classrooms the effects of reverberation on cognitive performance, annoyance due to noise and social-emotional attitudes to school was measured.

Learners that attended classrooms with favourable RT were compared to those with less favourable reverberant rooms. The results from the study indicated that performance was poorer in classrooms with longer RT than those with shorter RT. This left the authors to conclude that permanent exposure to unfavourable acoustic conditions in the learning environment may impair the development of auditory functions which are relevant in learning to read and spell (Klatte et al., 2010a).

Similar findings were reported by Ljung et al. (2011) who studied the effect of reverberation on memory recall of spoken lectures. Twenty adolescents from the upper secondary school participated, 19 had normal hearing thresholds (17 female and two male). The results demonstrated that when the lectures were presented in a room with long RT (1.84s at 125Hz, 1.46s at 250Hz, 0.94s at 500Hz, 0.77s at 1kHz, 0.78s at 2kHz and 0.68s at 4 kHz) the memory performance was poorer than the results in rooms with short RT (0.3 seconds, frequency range 125Hz to 4kHz).

1.1.4 Issue 2 – the curriculum and noise

Education systems both nationally and internationally have evolved, with traditional teaching methods supplemented by more informal, learner-centred and interactive teaching approaches (Fisher, 2007, Heeney, 2007). The Scottish Government in 2010 implemented a new national curriculum, The Curriculum for Excellence (CfE) covering learners from the age of 3 to 18. There are eight curricular areas: expressive arts, health and well-being, languages, mathematics, religious and moral education, sciences, social sciences, and technologies. Many of the teaching methodologies associated with the new curriculum: collaborative working, group work, active learning, and independent research, are generally associated with increased levels of movement and speech-based noise (Fisher, 2007).

Furthermore, the curriculum requires young learners to engage in a range of complex listening activities which can require attention to be divided between multiple talkers who may not always be visible (Valente et al., 2012). Accessing spoken communication generally involves the learner attending to, evaluating, comprehending, and reacting to the target signal of interest (voice of the teacher or learner) in a complex process that incorporates simultaneous sensory, linguistic and cognitive functions (Hällgren, 2005, Rönnberg et al., 2014). In suboptimal conditions, it can be difficult for the learner to fully access the target signal of interest and this requires increased listening effort. This process is in part determined by the linguistic and cognitive capacity of the learner, the quality of spoken instruction and concepts being taught (Yee Choi and McPherson, 2005).

1.1.5 Issue 3 – young people and listening in noise

Age is a significant predictor for reduced speech perception in noise. It is acknowledged that the aging process can result in adults finding it difficult to follow conversations in noise; this is partly due to declining auditory, cognitive and central processing functions (Schneider et al., 2005, Anderson et al., 2013). In contrast, the adverse effects of noise on listening for young learners are primarily due to maturation and experience. Young people have an immature auditory system that is going through a process of development and refinement which affects the ability of the young learner to use spatial cues to segregate the target signal of interest from noise (Boothroyd, 1997, Klatte et al., 2013). Furthermore, young learners do not always have a fully developed knowledge of linguistic rules and life experience to compensate for the effects of a degraded target signal of interest (Elliott, 1979).

1.1.6 Issue 4 – the attainment gap

Learners living in areas of deprivation have historically maintained lower levels of educational attainment compared to the general population (O'Neill et al., 2014, Sosu and Ellis, 2014). In Scotland, one in five children live in poverty and there is a gap in educational outcomes between learners living in the most and least deprived areas of the country. This is commonly referred to as the attainment gap and is measured by the proportion of young people from the least and most deprived areas that are performing well or very well at their expected curricular level (Scottish Government, 2016c). Closing of the attainment gap is a priority for national and local governments. In reading, the attainment gap at the primary four stage was 18 per cent between those from the most and least deprived areas. There was a 15 per cent gap in writing and 14 per cent in listening and talking (Scottish Government, 2016c). A similar pattern was observed in numeracy with a 21 per cent gap between those from the most and least deprived areas (Scottish Government, 2015b). There is also a focus on improving educational outcomes for all young people with between 70 to 80 per cent of learners in Scotland only achieving the expected level in numeracy, writing, reading and listening and talking (Scottish Government, 2018a).

1.1.7 Intervention – Dynamic soundfield

Sensorineural deafness is commonly associated with difficulties of listening in noise. Plomp (1978) regards deafness as a spoken communicative impairment, primarily in noisy environments. Deafness for speech is characterised by attenuation (reduced levels of both speech signal and noise) and distortion (decrease in the signal-to-noise ratio). The signal refers to the target speech signal of interest, this could be the teacher or learner and the noise refers to any masker that interferes with the signal being heard. The relationship between the two is commonly referred to as the signal-to-noise ratio (SNR) (Gelfand, 2016). For students with sensorineural deafness, the distortion factor is the most significant concern in the classroom as a hearing aid can compensate for the attenuation but not the distortion (Plomp, 1978, Dillon, 2001). To mitigate the reduced levels of speech intelligibility in noise either personal radio aid systems or whole class soundfield technology have been recommended for use with deaf learners.

1.1.8 Definition of a dynamic soundfield system

A soundfield system is a device that provides mild amplification to the voice of the teacher. It generally consists of one or more of the following items: transmitter, speaker and receiver/amplifier. There are both fixed and portable soundfield systems. A fixed system has wall or ceiling mounted speakers that are linked to an amplifier/receiver which connects to the transmitter worn by the teacher. A portable system has a similar configuration but instead of fixed speakers located across the room, the speakers are encased in a portable unit which means it can be transported between rooms. There are three commonly used transmission signals: infrared, frequency modulation and digital 2.4GHz.

The current study fitted the Digimaster 5000 dynamic soundfield system which operates on the 2.4GHz platform. The system automatically monitors background noise and adapts the gain level of the teachers' voice. This is referred to as a dynamic soundfield system (Dance et al., 2018). When classroom noise levels are <54 dB SPL, the gain is kept at a value of +6 dB. The dynamic adaption is applied at noise levels >54 dB SPL. Noise levels of 55 dB SPL in a classroom achieve a signal-to-noise ratio of >12dB, noise levels of 65 dB SPL have a signal-to-noise ratio of >10dB (Phonak, 2014). Soundfield systems have been widely used in educational settings to support inclusion and to improve educational outcomes of deaf learners in mainstream schools and the specialist type provision. The purpose of the assistive technology is to increase the level of the target signal of interest relative to the background noise, reduce listening effort, increase speech intelligibility and deliver improvements in attainment (Maltby, 2002).

Although there have been several studies into the effectiveness of soundfield there is very little evidence to support the use of soundfield technology with young hearing learners, and specifically those living in areas of social deprivation. Previous research has been compromised by poor study designs that were susceptible to response bias, recruitment of small sample sizes, unpublished papers and articles that were not peer reviewed and did not incorporate standardised educational assessments that were blindly marked (Stephenson, 2007). To date, there has been no longitudinal research into the effectiveness of dynamic soundfield on educational attainment.

Therefore, it appears timely to consider the effectiveness of dynamic soundfield technology on delivering improved educational outcomes and reducing the poverty associated attainment gap with young hearing learners. The current research aims to quantitively assess whether dynamic soundfield technology improves educational attainment for young hearing learners (age 7-8) in mainstream primary schools. Furthermore, the study will assess whether dynamic soundfield technology has a greater effect on learners from the most deprived quintiles compared to those from the least deprived quintiles.

1.2 Motivation

The motivation for the study stems from my professional role as an educational audiologist supporting deaf learners. Many young people attend our joint education/health clinics who experience challenges with listening in noise, even though they have no permanent form of deafness. In general, young people attend schools where the overall suitability of the school building and acoustic conditions is either poor or unknown. At the start of the study in 2012, 20 per cent of the primary school estate in Scotland had a suitability level that was categorised as poor. In 2018 that figure had reduced by only three percent to 17 per cent (Scottish Government, 2018c). There is no published data on the acoustic properties of the Scottish school estate, although it is known that older buildings generally do not meet government recommendations on ambient noise levels and T_{mf} (Shield and Dockrell, 2004). If the conditions of a school are categorised as poor, it can be inferred that the acoustic environment is also unsatisfactory.

The Scottish Government's committed 1.8 billion pounds to construct or refurbish 117 new primary and secondary schools by March 2020 (Scottish Government, 2018b). Local authorities have the discretion on the priorities for capital expenditure on school estates and Fife Council is committed to a construction and refurbishment programme under the name Building Fife's Future (Fife Council, 2009). As the educational audiologist, my professional role was to advise on the inclusive technology for the new schools. Dynamic soundfield was a core component of the Council's access strategy for deaf young people and adults and there was a discussion around how inclusive technology could contribute to one of its other objectives: improving educational attainment. During these discussions, the Head of Education requested empirical data on the effectiveness of dynamic soundfield on the achievement of young people living in areas of social deprivation. This research was a response to a local issue driven by national priorities.

1.3 Contribution to the field

This thesis is concerned with the twin aims of the Scottish Government's education reforms: closing the poverty-associated attainment gap and improving educational attainment for all. To deliver improvements in equity and achievement the Scottish

Government introduced the National Improvement Framework which set priorities for improving literacy and numeracy, health and well-being, positive school leaver destinations and closing the gap between the most and least deprived in society (Scottish Government, 2018a). As part of this policy in 2018, the Scottish Government introduced the Scottish National Standardised Assessments which is a web-based adaptive assessment programme administered to all learners at primary stages 1, 4, and 7 and secondary level 3. The assessments allow educational progress to be measured and monitored at a national and local level (Scottish Government, 2018c).

Before the national assessments, the majority of Scotland's local authorities used a similar web-based assessment and monitoring system called the Achievement for Excellence (Interactive Computerised Assessment System) (AfE (InCAS)). The assessments were administered by the Centre for Evaluating and Monitoring (CEM) at the University of Durham. The current research used the CEM assessment to evaluate the effectiveness of the dynamic soundfield system and all the assessments were marked by the University of Durham who were blind to the intervention. To date, there has been a paucity of quantitative research into educational outcomes of learners from areas of social deprivation using national standardised measures of academic achievement. This research will significantly contribute to the knowledge of policymakers, practitioners and government by providing quantitative data on the effectiveness of using classroom amplification to improve educational attainment, specifically for learners living in the 20 per cent most deprived areas of Scotland.

As part of the Scottish Government's improvement strategy, a Pupil Equity Fund was established to provide headteachers with direct funding to close the poverty-related attainment gap. In addition, Regional Improvement Collaboratives were required to produce improvement plans using robust evidence. To date, there has been no longitudinal research into the effects of dynamic soundfield technology on the educational attainment of learners in the least and most deprived quintiles using standardised national assessments. This research will allow Headteachers to make informed decisions on resource allocation of the Pupil Equity Fund.

To date, there has been a paucity of noise and acoustic surveys into Scottish schools since the new CfE was officially introduced in 2010. This research will develop an understanding of the noise levels during different lessons and learning activities. This research will provide valuable information to policymakers, politicians and those making decisions on new school buildings on the use of technology to improve educational outcomes.

1.4 Organisation of the thesis

This thesis is divided into eight chapters. Chapter 2 introduces and reviews the literature and concepts that are relevant to this study. Chapter 3 describes the methodology and study design including recruitment, equipment, and ethics. Chapter 4 presents the data on the participants, teachers, and schools to demonstrate both external and internal validity. This chapter also presents the descriptive and inferential statistics on room acoustics and the use of the dynamic soundfield system. Chapter 5 provides the results from the noise surveys completed in the control and intervention classrooms. Analysis of the learner and teacher questionnaire is also presented. Chapter 6 presents the results from the pre and post-intervention AfE (InCAS) assessments. Chapter 7 provides the results from the AfE (InCAS) assessments of learners from the most and least deprived quintiles. Chapter 8 concludes the thesis by discussing the main findings of the study and the contribution this research has made to the field.

Chapter 2 Literature Review

2.1 Aims of the chapter

This chapter has two aims. First, to outline the relevant literature on the detrimental effects that noise has on speech communication for young learners with typical hearing thresholds. The second purpose is to provide a critical review of the literature on the use of soundfield systems in classrooms to better understand the limitations of previous studies.

Figure 1 provides a diagrammatic representation of the central themes of the first part of this review. Several conceptual models have been devised to define the contribution that the auditory system makes to communication and these generally contain four interactive components: hearing, listening, comprehension and reacting (Kiessling et al., 2003, Sweetow and Henderson-Sabes, 2004, Edwards, 2007, World Health Organization, 2007). Although based on the World Health Organization (2007) *International Classification of Functioning, Disability, and Health,* the models do not explicitly incorporate contextual factors, both personal and environmental which significantly influence the process of communication in everyday life. Classrooms are generally enclosed spaces that are acoustically complex as they contain multiple speech sounds that are partly a direct or indirect product of the occupants of the room. The seminal work of Cherry (1953) recognised the challenges presented when following speech in a multi-talker environment and identified a number of discriminating cues that could mitigate the effect. These included the spatial separation of the speech signal, speech reading and gestures, speech characteristics (gender, accents) and transitional probabilities where the listener uses prior knowledge and experience to predict the content and context.

Section 2.4 will explore these factors with reference to young people in the classroom. The chapter begins (Section 2.2) by providing an overview on the key functions of the peripheral auditory system on spoken communication. In Section 2.3 the psychoacoustic principles for listening in noise relevant to this research will be explored.

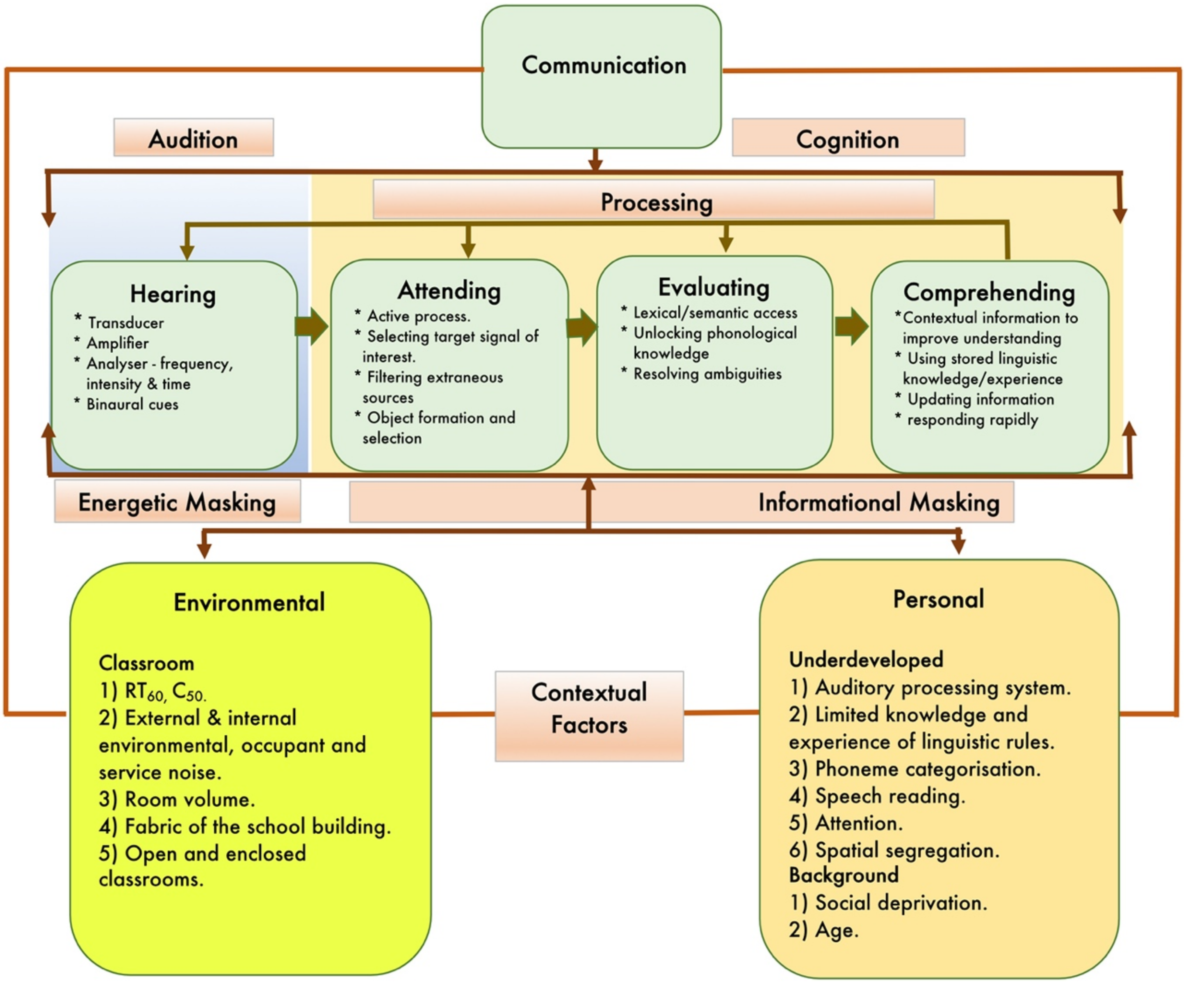


Figure 1: Diagrammatic representation of the processes associated with accessing spoken communication in a classroom.

Section 2.5 will critically review the literature on soundfield systems in the classroom and will highlight the paucity of research into the effectiveness of soundfield on educational performance for young people living in areas of social deprivation.

Section 2.6 of the chapter will conclude the literature review by highlighting the limitations of previous studies and outline the necessity for a longitudinal study into the effects of dynamic soundfield amplification on educational attainment that uses standardised national assessments that are blind marked.

2.2 Hearing – the peripheral auditory system

The peripheral auditory system transduces changes in air pressure produced by objects and events in the environment into neural responses (Boothroyd, 1997, Pichora-Fuller et al., 2007). Leibold et al. (2007) argue that to hear and obtain meaning from complex speech sounds the brain must be provided with an adequate sensory representation of the basic properties of sound. It requires to resolve the frequency, intensity and temporal characteristics of the original sound source.

The peripheral auditory system has three divisions: outer, middle and inner ear, see Figure 2. The primary function of the outer and middle sections of the ear is impedance matching as there is a mismatch caused by the transmission from one acoustic medium, air into another, the perilymph and endolymph composed fluid of the cochlea. Killion and Dallos (1979) computed the source impedance at the oval window by considering it as a small piston mounted in a baffle, the human head. The results from this theoretical model show that the impedance mismatch is frequency specific and may be greater than 50dB at 1000Hz. Therefore, the combined gain produced by the outer and middle ear needs to be considered to account for the loss of energy at the oval window interface.

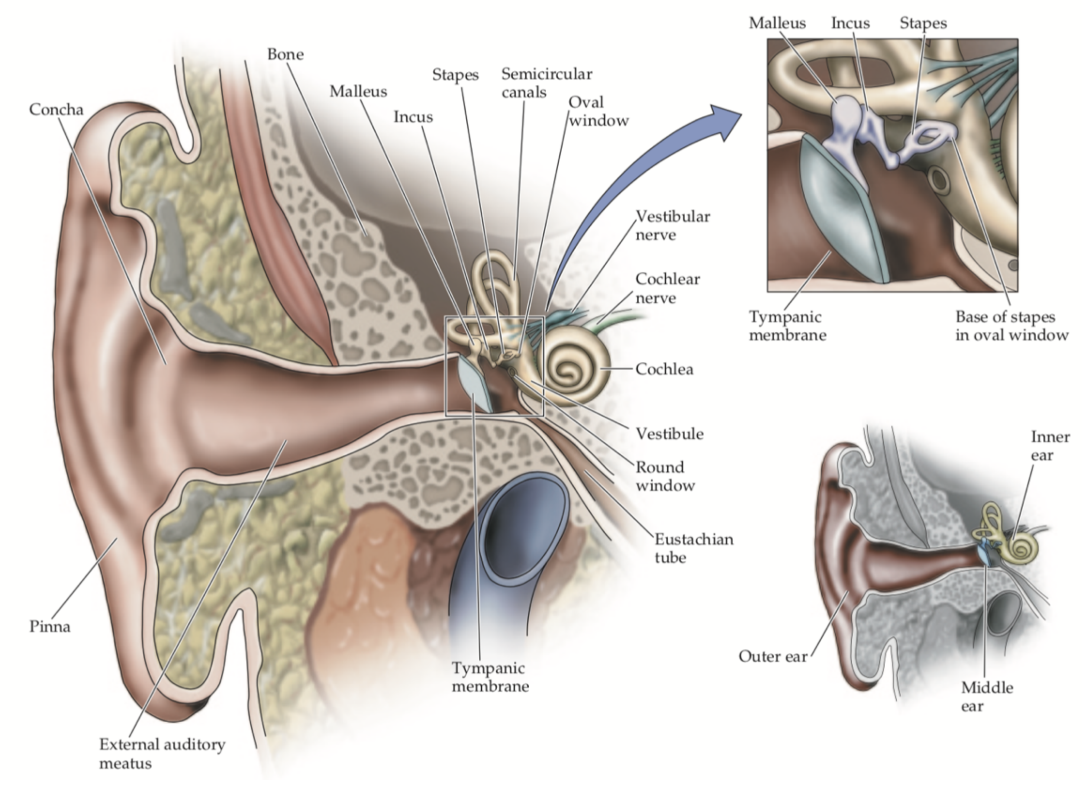


Figure 2: Schematic diagram of the peripheral auditory system (Bear et al., 2007).

2.2.1 Functions of the outer ear

The outer section of the ear consists of the pinna, concha, and external auditory meatus. Although the pinna has been associated with funnelling sound into the external auditory meatus, it has been demonstrated that hearing sensitivity is not affected when the pinna is excluded from sound conduction (Gelfand, 2016). The pinna's primary function is sound localisation, see section 2.3.4. The external auditory meatus is open at one end and closed by the tympanic membrane at the other, in essence, it is a quarter wavelength resonator. However, the external auditory meatus is an irregular shape and the length of the meatus varies with each individual which influences the resonant frequencies. Furthermore, the tympanic membrane and meatus walls are absorptive and so introduce dampening (Gelfand, 2016). Although there is individual variation, the common findings suggest that the resonant frequency of the external auditory meatus is between 2000-5000Hz (Shaw, 1974, Young, 2007). Mehrgardt and Mellert (1977) measured the transfer function of the external ear using twenty participants, with sounds presented in front of the listener at 0˚ azimuth; the gain provided by the external ear canal was between 10 and 15dB.

2.2.2 Functions of the middle ear

Three mechanisms of the middle ear also contribute to impedance matching: the area ratio of the tympanic membrane to the stapes footplate, the lever action of the ossicles and the curvature of the tympanic membrane. Gelfand (2016) applied the following formula to calculate the amount of energy that would be transferred from the air to the cochlea fluids without the middle ear, where T is the transmission and r is the ratio of impedance.

$$T = \frac{4r}{(r+1)^2}$$

The result is approximately 0.001, which suggests that 99.9% of the airborne energy would be reflected back rather than transferred. This represents a 40dB drop between the air and fluid.

2.2.3 Functions of the inner ear

The inner ear has three primary functions: transducer, amplifier, and analyser of sound. The cochlea contains three Scala (Gelfand, 2016). The Scala vestibuli and tympani are filled with perilymph, where the ionic composition is low in K^+ and high in Na^+. The Scala media is filled with endolymph, the ionic composition is ≈150mM K^+, 2mM Na^+, and 20μM Ca^{2+} and has a highly positive electrical potential of approximately +80 mV relative to blood plasma and perilymph. This is commonly referred to as the endocochlear potential and is essential for the transduction process, see Figure 3 (Hibino and Kurachi, 2006). As Figure 3 illustrates, the organ of Corti, sits on the basilar membrane and overhanging it is the tectorial membrane which is hinged at different points. Both the inner and outer hair cells are housed here and form synapses on neurons located in the spiral ganglion.

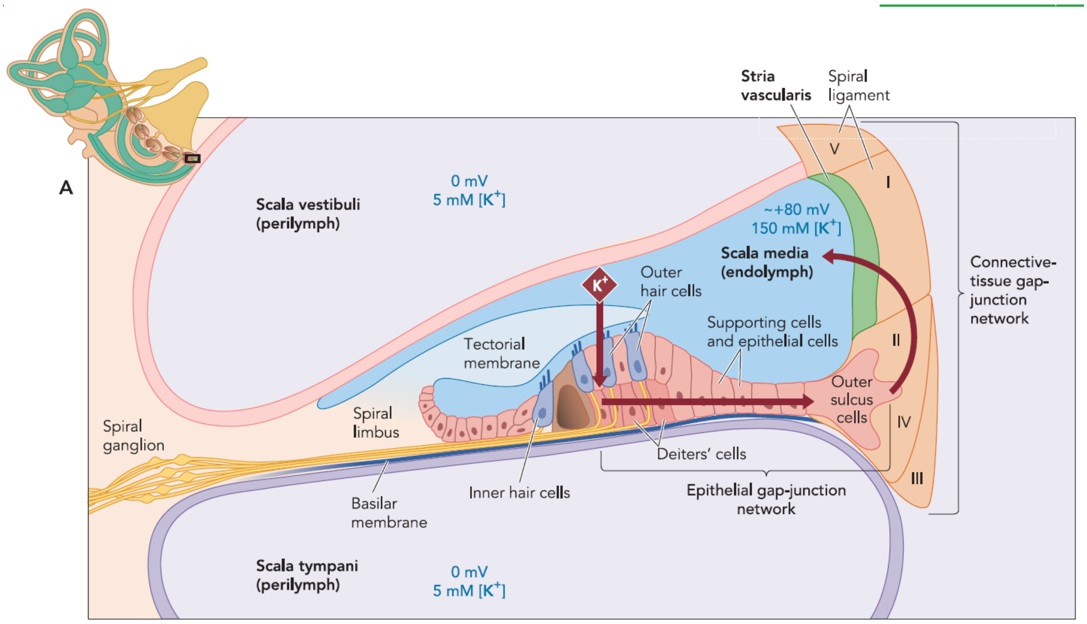


Figure 3: Schematic representation of the endocohlear potential of the inner ear (Hibino and Kurachi, 2006)

There are approximately thirty thousand neurons in the auditory nerve (Gelfand, 2016). Afferent nerve fibres carry information from the peripheral system to the brainstem. Ninety-five per cent of the afferent nerve fibres connect to the inner hair cells and so most information leaving the cochlea is from inner hair cells, see Figure 4. In contrast, outer hair cell innervation is predominantly efferent, carrying information from the brainstem to the cochlea with only a small number of afferent fibres (Frolenkov, 2006). The stereocilia sit on top of the sensory cells and are of different lengths, arranged radially and contain an actin cytoskeleton which makes them stiff. Furthermore, they are joined by tip links and so move as a unit (Gelfand, 1998). The stereocilia of the outer hair cells are attached to the tectorial membrane whilst the inner hair cells are not. Sound vibrations entering the cochlea create a shearing motion between the basilar membrane and tectorial membrane and cause the deflection of the stereocilia (Bear et al., 2007, Guinan et al., 2012). The deflection of the stereocilia opens mechanoelectrical transduction channels and allows the entry of K^+ from the endolymph that depolarises the sensory cells which then triggers the inflow of Ca^{2+} and leads to the release of neurotransmitters from the sensory cells to the associated neuron (Moore, 2003, Bear et al., 2007).

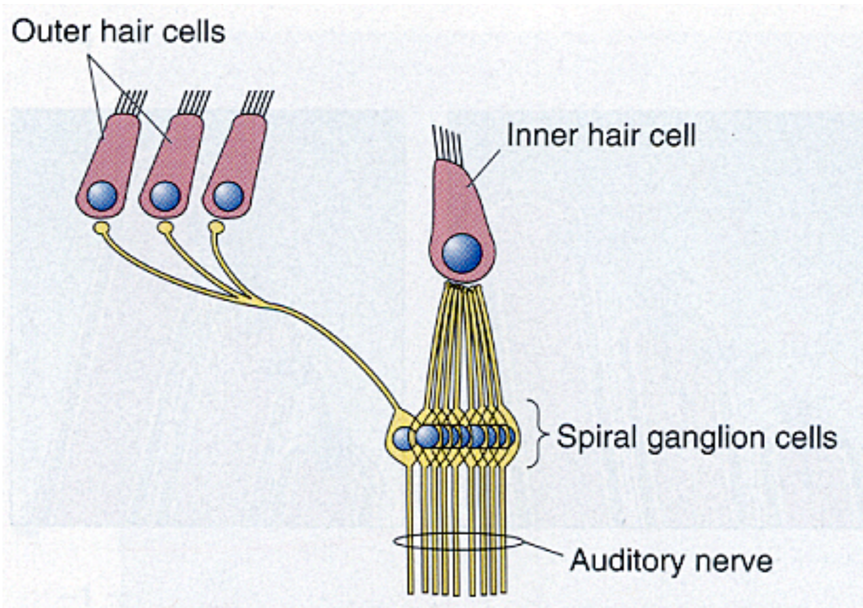


Figure 4: Inner hair cells are predominantly inverted by afferent fibres and outer hair cells mainly by efferent fibres (Bear, 1996).

The opening and closing of the mechanoelectrical transduction channels lead to the outer hair cells contracting and elongating when the cells are depolarised and hyperpolarised, see Figure 5. These conformational changes in the outer hair cells increase the movement of the basilar membrane and is one reason they are often referred to as the cochlea amplifier (Ulfendahl and Flock, 1998, Bear et al., 2007). Gummer and Preyer (1997) observe that just as the middle ear provides matching for the low impedance motion of air particles to the high impedance motion of cochlea fluids, the outer hair cells perform a similar function to compensate for the impedance mismatch between the perilymph and endolymph fluids, cochlea partition and the stiffness of the stereocilia. Moore (2003) describes the compressed non-linearity of the hearing mechanism. For low and medium sound levels, the gain from the outer hair cells was approximately 50dB. For low inputs, the gain is roughly constant and as the sound level increases the gain reduces and at 90dB SPL there is no gain contributed by the outer hair cells. Outer hair cells play a significant role in the amplification, sensitivity, and selectivity of human hearing.

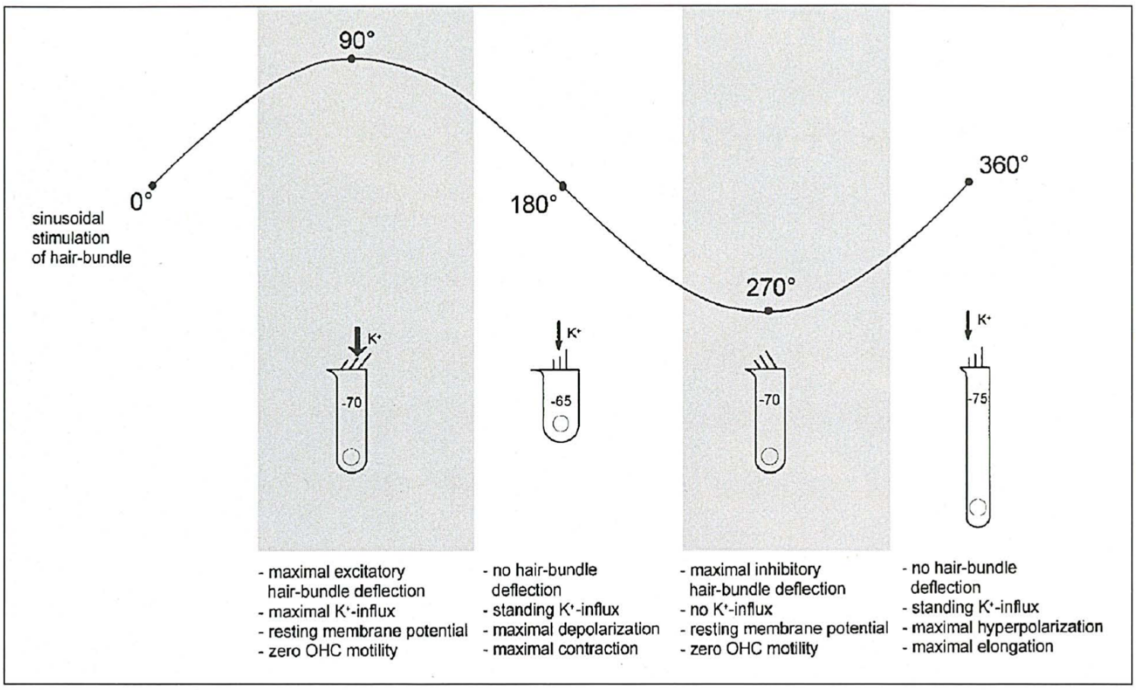


Figure 5: Contraction and elongation of the outer hair cells during one stimulus cycle (Gummer and Preyer, 1997).

The conformational changes in the outer hair cells are perceived by the inner hair cells which convert the movement of the basilar membrane into action potentials or spikes and the discharging of a spike is called firing, see Figure 6. The number of action potentials discharged per second is called the firing rate (Gelfand, 2016). The more intense the original sound source, the larger the area of basilar membrane vibrates. Furthermore, the greater the vibration, the greater the depolarisation and hyperpolarisation which in turn increases the spike rate.

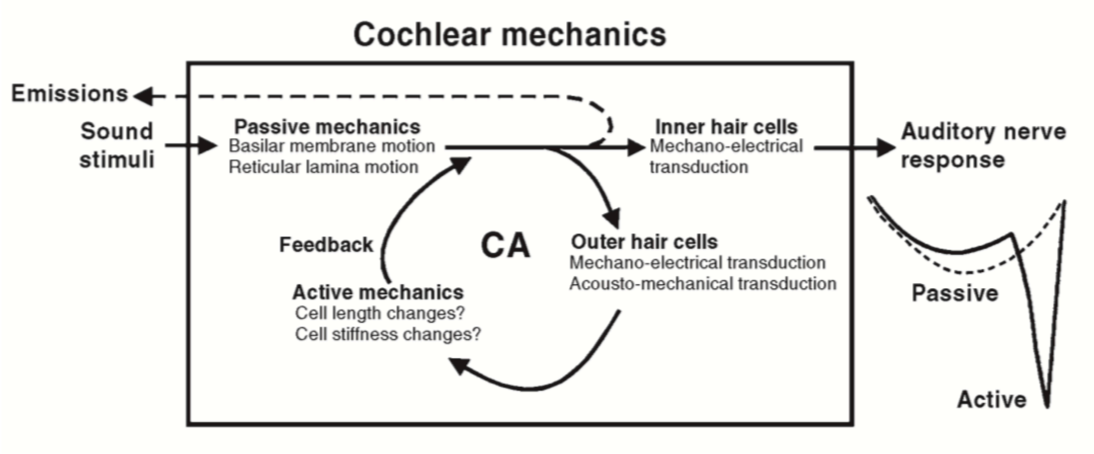


Figure 6: Schematic diagram illustrating the role of the outer hair cells in cochlear mechanics (Ulfendahl and Flock, 1998).

The inner ear also performs the function of frequency analysis, which is a central component of speech processing. There are two main theories of frequency analysis: place and timing theories. The tonotopic or frequency-to-place organisation of the auditory system was established by Bekesy's physiological experiments on human cochlea in the 1960s (Coleman, 1961). Sound pressure generates a travelling wave of the same frequency that moves away from the stapes, along the length of the basilar membrane towards the apex, see Figure 7. The amplitude of the travelling wave grows and velocity slows reaching a maximum at its resonant characteristic place (Bear et al., 2007). As Figure 7 illustrates, the place-frequency map of the basilar membrane sees high frequency sounds generating responses at the base and low frequency sounds at the apex. As speech sounds are complex waveforms containing multiple frequencies then multiple places on the basilar membrane vibrate. The cochlea is said to act as a frequency analyser as it decomposes the original speech signal into simpler components (Moore, 2008).

Bekesy's experiments were undertaken on passive human cochleae taken from cadavers (Manley et al., 2012). Over recent years the importance of the active role played by outer hair cells in frequency sensitivity and selectivity has been established (Moore, 2003). Research originally undertaken on animals has suggested that the tonotopic organisation of the cochlear is preserved throughout the central auditory pathways (Young 2007).

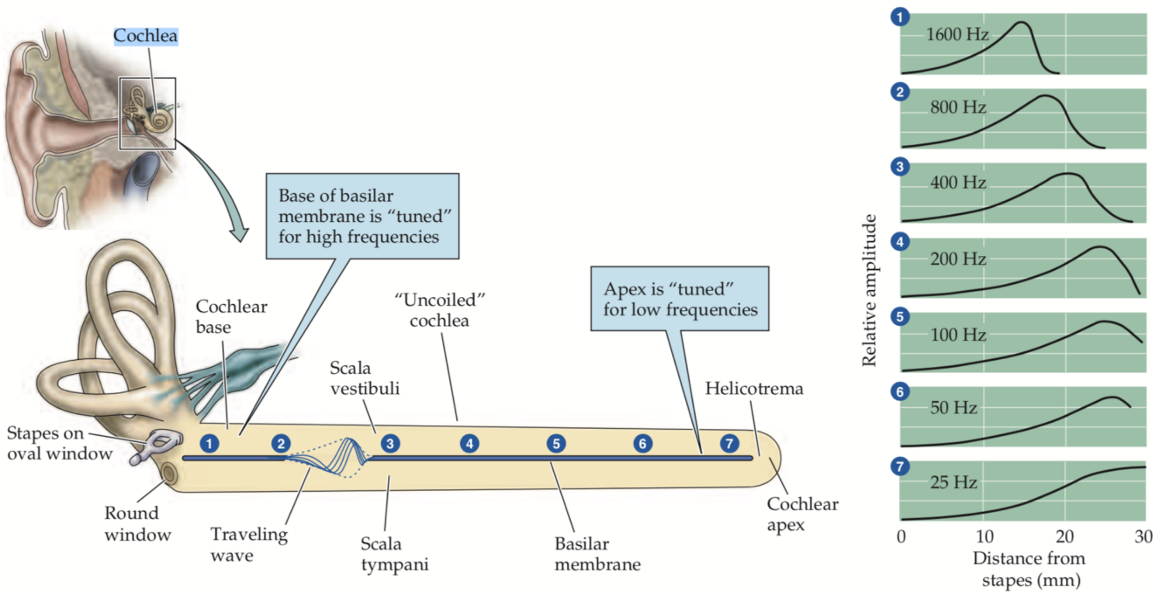


Figure 7: Schematic diagram of the tonotopic mapping of the inner ear (Bear et al., 2007).

The second frequency coding strategy is known as temporal theory (Bear et al., 2007). The vibration of the basilar membrane results in the displacement of inner hair cell stereocilia but the hair cells only release transmitters when depolarised and therefore action potentials are only discharged during the positive phase at low frequency. This is known as phase locking as firing only occurs at a particular phase in the signal which is synchronised to the frequency of vibration of the place on the basilar membrane. Plack (2012) observes that frequency is coded by both place (the location on the basilar membrane where the neurons are active) and phase locking (the temporal regularity in the pattern of firing). Although there is no specific cut off point, it is generally acknowledged that both frequency-to-place and phase locking functions below 5000Hz. Above 5000Hz frequency is coded by place only (Bear et al., 2007, Plack, 2012, Gelfand, 2016).

2.2.4 Summary - the function of the peripheral auditory system

Objects and events create changes in air pressure that are amplified at certain frequencies by the outer and middle ear. The inner ear decomposes the signal into simpler components, with the result that the frequency, intensity, and phase of the original signal is coded by neural responses in the auditory nerve. The place, timing, and rate of the firing of neural impulses is significant for the encoding of the frequency, time and intensity of the original signal. Moore (2003) rightly observes that all sounds are subject to this analysis within the cochlea. This tonotopic representation is maintained in the ascending auditory regions of the brainstem (including the auditory cortex and temporal lobe) and is significant for speech processing. The perception of these speech sounds as a coherent whole depends upon the representation of the individual decomposed components being reassembled at a later stage in the central auditory system.

2.3 Psychoacoustic basis for hearing speech in noise

The limitations of the auditory system to resolve speech sounds in background noise has been demonstrated by psychoacoustic experiments into masking. Durlach (2006) observes that masking occurs when the target speech signal is degraded in the presence of another stimulus, the masker. The two main forms of masking and the ones of most interest here are energetic and informational masking.

2.3.1 Definition of energetic masking

Energetic masking generally refers to physical masking that occurs within the peripheral auditory system and much of the knowledge was a product of the band-widening experiments performed by Fletcher in the early 20^{th} Century. Fletcher (1940) suggests that the peripheral auditory system behaves like a bank of overlapping bandpass filters. As revealed in Section 2.2 each place on the basilar membrane responds to a certain range of frequencies with each point having an auditory filter that corresponds to a different characteristic frequency. Moore (2008) observes that three assumptions are made when listening to a sinusoidal signal in the presence of broadband noise: the listener makes use of an auditory filter with a centre frequency close to the signal and the filter removes most of the noise. Only noise passing through the filter will mask the signal. Finally, the signal to noise ratio is determined by the amount of noise passing through the filter. The strongest masking is thought to occur at the centre frequency and the level of masking increases as the intensity of the masker is raised. Although the amount of masking increases as the noise passes through the auditory filter, once the noise bandwidth exceeds the filter bandwidth further increases in noise does not increase the level of masking (Moore, 2003). Energetic masking can result in portions of the speech signal becoming inaudible or degraded at the periphery if the simultaneous speech signals contain energy in the same critical bands at the same time (Brungart, 2001). Environmental noise generally produces energetic masking of speech whilst speech-based noise results in both energetic and informational masking (Arbogast et al., 2002, Schneider et al., 2007). In contrast to energetic masking that is associated with a degraded representation of the auditory signal when action potentials are discharged, informational masking is associated with sufficient representation of the auditory signal in the cochlea and beyond (Arbogast et al., 2002).

2.3.2 Definition of informational masking

The literature on informational masking provides a variety of explanations and causes (Cooke et al., 2008, Kidd Jr et al., 2008). Informational Masking is generally associated with perceptual masking that occurs within the higher level of the auditory system. Durlach (2006) argues that defining informational masking as everything that reduces speech intelligibility in noise that is not energetic masking fails to address the complexity of the issue. It is not a simple binary position of energetic (peripheral) versus informational (central) as there is also energetic masking in some of the higher central channels. Adopting a broad-brush approach was also criticised by Watson (2005) who argued that informational masking is a term that has become a suitcase word, in which a number of loosely related, but potentially different ideas are packed together.

Instead of ambiguous definitions of informational masking, it is more productive to analyse the effects. Informational masking may be a consequence of similarities between the target signal of interest and interferer, or uncertainty over which speech source is the target signal of interest and which is the interferer (Arbogast et al., 2002). Furthermore, parts of the auditory signal of interest and competing speech masker may be misallocated which can cause the listener to receive inaccurate information. Resolving this misallocation involves the listener expending greater amounts of listening effort (Cooke et al., 2008). Informational masking affects both object formation and object selection. Object formation influences how we perceive complex environments by focusing on one object (target signal of interest) and disregarding others (speech masker). Object selection involves the use of cognitive resources to attend to the correct target signal of interest (Shinn-Cunningham, 2008). In a classroom, the learner can be engaged in tasks that require following a discussion involving multiple talkers and so there is a need to attend to and track a different target signal of interest over time. Where there are similarities between the target signal of interest and masker then both object formation and object selection can be compromised. Informational masking effects speech perception, listening effort, and attention.

2.3.3 Binaural unmasking and better ear listening

It is generally acknowledged that where the target speech signal of interest and background speech sounds are separated on the azimuthal horizontal plane there is improved performance in noise (Gelfand, 2016). This is referred to as the spatial release from masking or unmasking of speech (Brungart and Simpson, 2002, Dubno et al., 2008, Viswanathan et al., 2016). This is primarily due to two different strategies used by the listener to localise sound and effectively listen in noise: interaural time difference (ITD) and interaural level difference (ILD). These are frequency and location specific. Any sound directly in front of the listener will be received by both ears at the same time and so there will be no ITD. Conversely, a sound source at one side of the listener (90° azimuth) will take longer to reach the ear furthest from the source, resulting in an ITD. If the wavelength is equal or greater than the width of the head, then it will diffract around it. Due to the different wavelengths of high and low-frequency sounds, ITD is only applicable to low-frequency sounds (Gelfand, 2016). Furthermore, the distance between the two ears means that the sound wave arrives at each ear at a different phase in the wave cycle and so these differences in the ITD perceptually separate the target from the competing noise (Moore, 2007). For higher-frequency sounds the localisation of sounds depends on ILD. This is often referred to as the head shadow effect as high frequency sounds are attenuated by reflections off the head. Consequently, high frequency sounds at the ear closest to the sound source will be more intense compared to the other ear. ILD is also affected by the sound level difference at the two ears caused by the different distance to the source. ILD is generally associated with improvements in audibility at the ear that receives the higher target to masker ratio due to the head show effect (Edmonds and Culling, 2006). Srinivasan et al. (2016) observe that primarily the spatial release from masking is a combination of binaural cues (ITD) and better ear listening (ILD).

2.3.4 Summary of masking on speech

When spoken communication occurs in a complex acoustic environment, two types of masking can compromise intelligibility: energetic and informational. Energetic masking is commonly associated with the physical degrading of the target signal of interest at the peripheral level. Informational masking is interference that occurs at the perceptual level and commonly stems from uncertainty in the stimulus (Arbogast et al., 2002). Masking can be mitigated through a variety of cues including the spatial separation of the target signal of interest and the interfering masking sounds. Binaural hearing means that each ear receives different signals that can be processed both independently and coincidentally, which facilitates the unmasking of speech.

2.4 Young people accessing speech in noise

2.4.1 Auditory Stream Segregation

Auditory attention is a key requirement of communication in a classroom situation as the listener must flexibly and selectively focus on the target signal of interest whilst simultaneously inhibiting other interfering sounds (Schneider et al., 2007). The ability to segregate the multiple overlapping signals into discrete auditory objects and identifiable events is referred to as auditory stream analysis. Bregman

(1993) suggests that two different processes are involved in separating sounds in the environment and identifying the source they came from. The first is a bottom-up process that is innate to all animals and involves the partitioning a complex mixture of sounds into separate acoustic sources using the general acoustic cues that are present in virtually all environments. The second involves a top-down process that requires conscious attention which has developed through knowledge and experience.

Cherry (1953) believed that the separation of the sound sources would assist with the unmasking of sounds in a multi-talker environment. A number of studies have demonstrated that young people have immature auditory attentional skills compared to adults. For example, Johnstone and Litovsky (2006) compared the effect of three different maskers (speech, reversed speech and modulated white noise) had on a group of 20 young people (aged between 5-7) and 20 adults (aged between 18-42). The study found that regardless of the type of noise, location of noise and the difficulty of the task, the speech recognition threshold for young people was always higher than that of adults. Young people were more susceptible to the effects of masking, especially when spatial cues were removed. Furthermore, it would appear that young people are more likely to stream both relevant and irrelevant acoustic streams and have difficulty selecting the target signal of interest and maintaining focus over time.

Speech is a complex sound containing acoustic segments of energy such as plosives, fricatives, and formant transitions and as such without directing attention and filtering then misallocation can occur for any part of the speech sound (Cooke et al., 2008). The auditory components of speech sounds are generally measured in three parameters: intensity, time, and frequency. Sussman et al. (2007) measured the frequency separation thresholds, the point at which a complex sound is perceived as being streamed as one stream instead of two streams in 40 school-aged learners (ages 5-8 and 9-11) and 14 adults (ages 19-24). It found that stream segregation by frequency separation was subject to maturation. With increasing age, a smaller frequency separation was required between the two sounds to hear them as two distinct streams. The author concluded that higher level processes within the auditory system required experience to achieve adult-like levels of acuity.

Perceptual auditory grouping and object formation theory evolved from the work of Bregman (1993). Although the acoustic environment is a complex mixture of overlapping signals, the ability of the listener to identify different objects is possible as groups of sounds are simultaneously grouped together. This includes location, onset time, harmonicity, timbre and pitch (Darwin, 1997, Best et al., 2007). If the object is distinct or the listener knows the features of speech, then the correct target is selected (Bregman, 1993, Shinn-Cunningham and Wang, 2008). This is in keeping with other studies that have demonstrated prior knowledge of the talkers' voice, the topic being discussed, and spatial location of the target assists the listening process (Brungart, 2001, Kidd Jr et al., 2008, Schneider et al., 2007). However, informational masking can result in the wrong individual object being identified if the listener does not know what acoustic features to attend to. This theory is supported by Brungart and Simpson (2004) research into masker uncertainty which measured the effects of informational masking. The test materials were the Coordinate Response Measure (CRM) which is a seven-word matrix sentence in which a noun is selected from a group of eight. The colour and numbers in the sentences were selected at random from a set. The masker was the same sentence structure but with different nouns, colours and numbers. The results showed that the majority of incorrect responses originated from the content presented from the masker. The researcher concluded that when the listener is able to hear two simultaneous phrases in the same ear, they are generally able to identify the words (object) but not which talker (object selection) they are coming from.

Wightman and Kistler (2005) assessed the effects of age and gender on the release from masking using the CRM materials on 38 young people (4-16 years) and eight adults. Due to the number of variables in such a broad age range, the young people were assigned age groups spanning approximately two years. The procedure involved presenting in the right ear a simultaneous target and distractor phrase. The target signal was always the voice of a male and the distractor was sometimes a male and sometimes a female. In the contralateral ear, there were three different masking conditions: no presentation, speech-shaped noise, and CRM phrase. For both adults and young people, there was a substantial release from informational masking when the target talker was male and distractor female. The opposite-gender effect on the release from masking is replicated in other studies. Brungart (2001) used eight target signals (four male and four female) and a distractor composed of male and female voices. The results demonstrated a significant improvement in performance when the target signal and masker were from different genders. These results support the Cherry (1953) hypothesis that gender differences in speech characteristics can spectrally provide a release from masking. The results also showed a significant age effect, in which even the oldest group of children (13.6-16 years) demonstrated poorer performance than adults at establishing object selection. The majority of errors between the target signal of interest and distractor were in the younger group of children.

Leibold et al. (2007) explain that young people listen less selectively than adults and so do not focus on the important features of speech but stream the whole sound. This broadband approach is the consequence of age-related developmental changes in neural structures and non-sensory processing. Young people will use different parts of the acoustic signal than adults to access speech and this goes through a period of development and refinement. Nittrouer (2002) termed the process as the Developmental Weighting Shift. Initially, the young listener will place greater emphasis on changes in formant transitions before giving greater weight to silent gaps, voicing duration, and place of consonant constriction.

2.4.2 Auditory stream segregation and social deprivation

There are a limited number of studies that explore the effects of selective attention and masking on young people that factored into the study design social deprivation. Jones et al. (2015) examined if selective attention explained the development of hearing speech in unpredictable noise. The procedure involved a tone detection test: the target signal was a 1 kHz pure tone and the unpredictable noise was 370ms long and presented simultaneously with the target signal. In keeping with the findings from the previous studies discussed, levels of performance in noise improved with age, achieving adult-like performance by 9-11 years. Multiple linear regression analysis showed that age was a significant predictor of masking, with masking decreasing by 1.8dB per year. This was associated with improvements in selective attention, with adults better able to ignore distracting noise similar in frequency to the target. The analysis also revealed that young people living in socially deprived areas were disproportionally affected by masking noise and this was attributed to factors such as working memory and concentration. Interference from noise was more of a distraction when the task required a greater consumption of cognitive resources.

2.4.3 Speech reading and unmasking of speech sounds

Cherry (1953) identified speech reading and gestures as other factors that would help unmask sounds in a multi-talker setting. The benefits of concurrent auditory and visual information when the speech signal is masked has been demonstrated in a number of studies. Helfer and Freyman (2005) compared the effects of speech reading on the ability to hear a single female speaker using two different interferers: steady-state masking noise and a complex masker composed of two female voices. The experiment was completed with and without spatial separation. The results showed that speechreading provided the greatest benefit when the masker was composed of speech compared to steady-state noise. Furthermore, visual cues were more effective when the target speech signal and masker were spatially coincident compared to when there was spatial separation. This suggests that visual cues are more effective at segregating the target signal of interest from the masker when the interferer is composed of other speech sounds and spatial separation is not present.

The majority of speechreading experiments have been undertaken with adult participants. There is limited research into the effectiveness of visual cues on the unmasking of speech for young people with typical hearing thresholds. Wightman et al. (2006) studied the effects that viewing the speaker had on informational masking. Twenty-three young people (6-16 years) and ten adults (18-31.9 years) with normal hearing thresholds participated. Three conditions were tested: audio-only, video-only and video and audio simultaneously. The results showed there was a strong relationship between audio-visual release from masking and speech reading ability, both of which improved with age. Young people required a higher SNR to achieve the same intelligibility levels as adults. Older young people and adults improved in the audio-visual condition, the release from informational masking for some adults was ≥15dB. However, for the youngest age group (6-8.9 years) the audio-visual condition produced no release from informational masking. The research found there was a correlation between audio-visual benefits and age (0.74) but when speechreading was factored out this reduced (0.03). This left the authors to conclude that this was primarily due to age-related changes in speechreading ability, which reflect age-related changes in the encoding of visual information in speech. Overall, the results suggest that young people are not as proficient at speechreading as older children and adults.

2.4.4 Language, phonemes and listening in noise

Johnson (2000) assessed 30 students with normal hearing thresholds (between 6 and 30 years) using a consonant-vowel-consonant-vowel phoneme identification test. Four different acoustic conditions were presented: quiet, reverberation only, noise only and a combination of noise and reverberation. In conditions of noise and reverberation adults were significantly better at identifying voiced or voiceless consonants than all the young people (ages between 6 years and 15 years, 11 months). Concerning the manner of articulation of consonants, a process that requires accessing rapid spectrum changes over time intervals of 10-30ms, the younger age group (6-7-year olds) were most affected. For the place of articulation, all age groups were significantly better than the younger age group. However, the mean score of the 10-11 age group was significantly lower than adults. Overall the results indicate that in combined noise and reverberation conditions it was not until the mid to late teenage years that maturation is achieved (Johnson, 2000). Young people require higher SNR to access more acoustical energy as they do not have the same perceptual skills in fine-grain discrimination and identification in the continua of synthetic speech items. Furthermore, noise and reverberation reduce the auditory cues available to the listener.

Further studies have also demonstrated the challenges that young people have with phoneme categorisation and discrimination, both in quiet and noise. Hazan and Barrett (2000) assessed the development of phoneme categorisation across a range of phonemic contrasts based on voice, place, and manner of production. The test was completed in quiet. Eighty-four young people (aged between 6 years and 12 years and six months) participated and the results indicated that young people’s phoneme categories are less well specified than adults and that consistency and boundary sharpening continues to mature into the second decade of life (Hazan and Barrett, 2000).

In addition, several studies have shown that young people, in general, do not have the same linguistic knowledge and phonemic discrimination skills as adults and are more susceptible to masking. Elliott (1979) deployed the Speech Perception in Noise (SPIN) assessment to study 24 students at four different age levels (17, 15, 13 and 11 years old) in addition to two groups of students at local schools in the 9 and 11-year-old age bracket. The SPIN assessment not only measures word intelligibility but the cognitive components of speech by comparing the results from high and low predictable sentences. The high predictable sentences contain two or three-pointer words that provide semantic clues to the keyword in comparison to the low predictable words that contain one pointer word with no semantic links. All participants had normal hearing thresholds and chronologically appropriate language ages. The noise was provided through a babble sound file with the SNR of +5dB, 0dB, and -5dB. Analysis of the results indicates that there is a differential for listening to sentences in noise related to age. Students aged 11-13 only performed more weakly than older students and adults in the highly predictable sentences whilst the 9-year-old students were significantly poorer in all the assessment in noise, and especially in the high predictable sentences, compared to the 11-year olds. The study concluded that the poor scores in the low predictable sentence assessment were primarily due to masking and the poor scores for high predictability was either the masking impaired the ability to use language rules or there was an immature knowledge of the rules amongst the young participants (Elliott, 1979).

Similar outcomes were observed when the speech signal was manipulated to remove spectral cues. Eisenberg et al. (2000) compared speech recognition in two groups of normal-hearing students (5–7 and 10–12 years of age) and one group of adults (between the ages of 21 and 55 years of age) listening to age-appropriate stimulus materials including sentences, words, nonsense syllables and digits in the same conditions. The test materials were manipulated using the noise-band processing technique based on the study by Shannon et al. (1995). This method involves the spectral information being removed from speech by the replacement of the frequency-specific information in a broad frequency region with a band-limited noise. The technique has the effect of preserving the temporal and amplitude cues in each frequency band, but the spectral information in each band is removed (Shannon et al., 1995). The scores between the older students and adults did not differ significantly, whereas the scores for the younger students were statistically poorer. The results indicate that young learners were not as skilled at using context as adults and do not have fully developed linguistic/cognitive capacity.

2.4.5 Language, phonemes, listening in noise and social deprivation

Cherry (1953) suggested that transitional probabilities including the content and context of speech could be used to overcome the barriers presented in suboptimal acoustic conditions. The assumption was the human brain has an enormous store of probabilities that could be exploited by the listener to predict information that may have been missed in noise. This infers a well-developed and sophisticated language system, including features such as phonological, lexical, syntactic and socio-linguistic. For some young people living in areas of social deprivation, this is not always the case. Studies have shown that young people living in poverty and children whose parents have no qualifications fared less well in literacy, language development, and social outcomes compared to children from higher socioeconomic classification and whose parents have higher educational qualifications (Bradshaw, 2011, Hartas, 2011). The widely cited study by Hart and Risley (1995) linked vocabulary learning, intergenerational poverty, and school failure. The 30 million word gap by the age of three is the most commonly cited statistic from the study (Kuchirko, 2019). Although criticised for lacking a nuanced portrayal of language interactions across different cultural and socioeconomic groups, there remains clear evidence of a persistent gap in language outcomes between young people from the most and least deprived backgrounds (Bradshaw, 2011, Sosu and Ellis, 2014, Hindman et al., 2016).

Nittrouer and Burton (2005) tested the hypothesis that early language experiences nurture the perceptual strategies required to access the phonemic structure in spoken language. Forty-nine young people aged (between 4 years 11 months and 5 years 11 months) were placed in one of four groups: a control, a group that experienced seven episodes of otitis media with effusion (OME) (a condition that impairs the impedance matching function of the middle ear, discussed in Section 2.2.2) by the age of three, low socioeconomic status with no recorded episodes of OME and a group containing both conditions. It was hypothesised that young people from backgrounds of social deprivation or those with transient deafness during the formative years would have impaired language experiences. The results showed that perceptual weighting strategies, syllable and phoneme awareness, word recall and comprehension of sentences with complex syntax structures were poorer. The only area in which the null hypothesis was not rejected was the temporal processing effects on language processing. The overall conclusion was that different perceptual strategies are required to extract phonetic structure from the target signal of interest. This requires a rich listening experience. Young people with a deficit in this area will find it challenging to use phonetic structure and consequently will have difficulty in storing and retrieving words from working memory and comprehending sentences with complex syntax.

2.4.6 Signal to Noise Ratio

Speech is a modulated signal both in time and frequency, and is both sparsely distributed at high energy regions and redundant (Cooke, 2006). Listeners can achieve a release from masking by taking advantage of the dips and glimpses in the target signal of interest when the SNR is improved. Buss et al. (2009) used the visual analogy of looking through a picket fence, where the gaps between the slats provide enough information to achieve an accurate picture of the scene beyond. A higher number of talkers in a classroom result in fewer glimpses. In a classroom where there are multiple talkers and/or long T_{mf} the effect is to decrease the number of dips and glimpses available in the target speech signal which affects speech intelligibility (Cooke, 2006).

Research has indicated that young learners require a higher SNR than adults. Bradley and Sato (2004) used the Word Intelligibility by Picture Identification test to measure speech intelligibility in noise in 41 classrooms involving 840 students in the 1st, 3rd and 6th grade with a nominal age of six, eight and eleven. Their findings indicated that younger people need a higher SNR to obtain the same speech intelligibility scores as older students. Grade 1 learners (aged six) required an SNR of more than 15.5 dB; grade 3 learners (eight years old) 12.5 dB and grade 6 learners (11 years old) 8.5 dB to obtain a 95% correct score rate. Similar findings were obtained by Vaillancourt et al. (2008) when establishing a Canadian French version of the Hearing in Noise Test. Seventy native-speaking French subjects with normal hearing participated, including 56 young people (aged 6-12 years) and 14 adults (18-30 years). A speech-shaped masker was used, and the results indicated that young people did not achieve adult-like performances in noise until the age of 12. The effect on performance was not limited to speech recognition thresholds but also affected spatial separation.

The correlation between noise and age has been supported in other studies. Blandy and Lutman (2005) measured the impact that speech recognition in noise had on young learners using the BKB sentence test. The research compared speech recognition test scores of 193 young people with typical hearing thresholds with 17 adults in a comparative study and found that young people aged seven required a higher SNR than adults. The young people in the study were between the age ranges of 7.1 to 8.3 years (mean 7.65) and the adults in the comparable study were aged 21 to 29. The assessment involved the BKB sentences in noise presented binaurally through headphones. An analysis of the results indicates that although 7-

year olds had hearing levels equivalent to or better than adults, their ability to listen to speech in noise was significantly poorer.

2.4.7 Cognition and language comprehension

It is recognised that working memory plays an active role in language comprehension and that there is increased cognitive effort when listening in conditions of noise and reverberation (Hällgren, 2005, Rönnberg et al., 2008). Listening to spoken communication involves not only detecting the speech signal of interest but applying meaning which requires attention, memory, language, and experience as we combine what we hear with what we know. Hällgren (2005) explains how understanding spoken communication involves both audition and cognition as signal driven and knowledge-based processes work in parallel. The learner makes use of sub-lexical information, perception, and identification of individual speech sounds and higher-order knowledge, including contextual, lexical, syntactic and semantic information processing to derive an accurate and comprehensive representation of the spoken information being communicated.

Kahneman (1973) proposed that working memory has a finite capacity. The more the listener relies on linguistic knowledge and past experiences to obtain meaning from speech, rather than bottom-up signal driven processes, the more cognitive effort is consumed. Valente et al. (2012) observe that in poor acoustic environments the increased effort to decode speech may leave fewer cognitive resources available for other activities such as language comprehension. Edwards (2007) argues that cognitive resources are expended during the process of communication itself. In classrooms, learners are listening and contributing during classroom discourse and during this process, the listener is reacting and creating relationships between different sentences while drawing higher-level contexts and storing information in both short-term and long-term memory.

The seminal work of Baddeley (2003) describes how working memory is divided into three sections. The first, the phonological loop is concerned with verbal and acoustic information whilst the visuospatial sketchpad is the visual equivalent. Both are slaves to the central executive which controls attention. The phonological loop stores linguistic information in a phonetic code. The listener makes use of perceptual strategies to access the phonetic structure from the target signal of interest. This is used to store and retrieve words from working memory which are required to comprehend sentences with a complex syntactic structure. Rönnberg et al. (2008) created a similar conceptual framework to describe the co-dependency between explicit and implicit cognitive functions and language comprehension called the Ease of Language Understanding (ELU) model. When the incoming linguistic signal relates to phonology, semantics, syntax, and prosody it is rapidly bound together at the cognitive level into phonological information. In optimal conditions, the linguistic input signal achieves rapid and implicit lexical access that is matched to the stored phonological representation in the long-term memory. Consequently, fewer cognitive resources are consumed. Conversely, when a mismatch occurs because of a degraded speech signal then explicit processing and storage capacity is required to infer meaning. This results in increased cognitive resources to correctly perceive the information being shared.

2.4.8 Cognition and divided attention

Increased listening effort is not only the consequence of mismatched or partial access to information but also a consequence of divided attention. Divided attention is commonly researched through the use of a dual-task paradigm in which two tasks are completed separately and concurrently. The primary task is generally a listening to speech in noise test and the secondary activity involves a more cognitively demanding task (Gosselin and Gagné, 2010). The overall purpose is to assess the effects that finite processing capacity has on each activity. This is relevant to the current study as it replicates common activities undertaken in a classroom such as multiple-talker discussions and taking notes/completing activities whilst the teacher is talking.

A study of 31 students between the ages of nine and twelve with normal hearing thresholds were tested using a primary task that involved repeating monosyllabic words from the AB Word List and the secondary task involved the storing and recalling a set of five digits presented on a computer screen. The noise babble was presented through multiple speakers with four different SNRs measured: quiet, +4dB, 0dB, and –4dB. The results indicated that performance reduced on the speech recognition and recall task as the SNR decreased (Howard et al., 2010). Choi et al. (2008) also used a similar method that required repeating monosyllabic words and digit recall. The results showed a decrease in performance in the secondary task and this was attributed to immature attention allocation. Overall the results from the study suggest that if learners are exerting more effort on the primary task of listening, this may limit the resources available in the secondary task and so compromise educational performance.

2.4.9 Summary and implications of listening to spoken communication in the classroom

Classrooms are complex auditory environments with multiple inputs masking the target signal of interest. Energetic masking and informational masking can compromise speech access. Unmasking of speech can be achieved through the use of several discriminating cues. Young people like adults can make use of the speech characteristics offered by opposite gender situations. However, spatial separation, speechreading, and other perceptual cues are subject to a process of developmental and experiential changes. Furthermore, the use of transitional probabilities to address issues of partial information may be compromised by finite knowledge and experience of language. Misallocation of resources and divided attention can increase the cognitive load that may compromise educational performance. Listening to spoken communication in the classroom involves the interaction and co-dependency of audition, language, and cognition and is relevant to the current research.

2.5 Research into Soundfield Systems

2.5.1 Methods and selection criteria

The literature on soundfields was searched using the following databases: DiscoverEd, Medline, PsycINFO, PubMed and Web of Science Core Collection. A search of the internet and journals was also performed. The inclusion criteria were studies that measured the effects of soundfield systems on academic performance or speech perception for young people in school settings. In addition, the participants mainly or wholly had to have typical hearing threshold levels. Excluded were small sample sizes (≤20), participants that were adults rather than young people and studies that had participants with congenital or acquired sensorineural deafness. The reason for excluding a small sample size was because this can result in bias and a loss of power; in general, auditory perception in adults is fully developed and in young people, it is going through a process of development and refinement and sensorineural deafness leads to broadening of the auditory filters and consequently is increasingly susceptible to energetic masking compared to the general population (Boothroyd, 1997). Eleven studies were identified that met the inclusion criteria. These have been divided into those that primarily used academic assessments to determine the effectiveness of the intervention, study designs that used questionnaires and those that used speech recognition assessments.

Table 1 provides the results and limitations of the eleven studies evaluating the effect of soundfield amplification on educational attainment and speech recognition. Slack and Draugalis (2001) observe that to establish that an intervention produced the outcome that is measured, extraneous factors that could have influenced the results must be discounted. In studies that examine the effects of amplification on listening the confounding variables include hearing acuity and the acoustic properties of the classrooms. Several of the studies did not include one or both of these factors. Furthermore, many of the papers were susceptible to response and expectation bias as the teachers were not blind to the rationale of the intervention. The study designs involved the teachers marking the learner's work, commenting on learner's behaviour and providing opinions on the effectiveness of the intervention.

STUDY	METHODOLOGY	FINDINGS	LIMITATIONS
MCSPORRAN ET AL. (1997)	**Design –** qualitative before and after study **Sample** – n=65, age range 7.3-8.2 **Soundfield type**- Phonic Ear Easy Listener **Acoustic Measurements -** RT was measured in both schools **Hearing Screening** – Pure tone audiometry (PTA) **Measures -** The Children's Auditory Processing Performance Scale (CHAPPS) questionnaire.	A significant number of the young people in the target group demonstrated an improvement in listening behaviour according to their CHAPPS scores.	The teachers rated the children's performance using CHAPPS and so not blind to the intervention. Potential response bias. No standardised measures of academic achievement. Lack of data on how often the soundfield system was used. No control groups.

STUDY	METHODOLOGY	FINDINGS	LIMITATIONS
ARNOLD AND CANNING (1999)	**Design** - Controlled clinical study. **Sample** – n=49 (n=23 female, n=26 male) primary school pupils from two upper un-streamed classes. Age range – 8.58-11.42. Mean age =9.92. **Soundfield type**- Frequency Modulation **Acoustic measurements** – occupied and unoccupied noise levels, RT recorded. **Hearing Screening** – not performed **Measures** - reading comprehension – Neale Analysis of Reading Ability (NARA). Student questionnaire	Significant improvement in comprehension over the three levels of the NARA in the rooms provided with soundfield.	Not a properly installed soundfield system. No hearing screening of the participants. Students had to write their answers to the questions and so the scores may have been affected by issues of literacy rather than listening. Participants and markers not blind to the condition.

STUDY	METHODOLOGY	FINDINGS	LIMITATIONS
ROSENBERG ET AL. (1999)	**Design** – two-phase longitudinal study. **Sample** –Phase 1 n=1319 kindergarten, first and second grade students. Phase 2 n=735 kindergarten, first and second grade students. **Soundfield type**- Phonic Ear Easy Listener four speaker system. Option of boom or lapel microphone. **Acoustic measurements** – ambient and occupied noise levels. RT not recorded. **Hearing Screening** – Phase 1- Pure tone audiometry (PTA) and tympanometry n=1252. Phase 2 – not performed.	The research found that in the classrooms fitted with soundfield systems the students significantly improved their listening and learning skills The greatest improvement was in the younger kindergarten age group.	Only some of the participants had a hearing assessment. There was no measurement of RT. The allocation of students to the control an intervention groups were not random. Teachers marked the assessments and they were not blind to the intervention and so response bias was possible. Training at the start of the research explained the

STUDY	METHODOLOGY	FINDINGS	LIMITATIONS
	Measures - Observational data was used to measure outcomes with the class teachers rating performance using the Listening and Learning Observation (LLO) Schedule. The mean difference in scores between the experimental and control group was compared.		benefits of the system and so introduced bias. Two different microphones were used in the trial which could affect the signal-to-noise ratio. Data on usage of the system was not provided. Mean test LLO scores were not statistically tested for differences between the control and intervention groups.
DARAI (2000)	**Design** – quasi-experimental longitudinal study	Greater literacy gains in the experimental classroom	There is no data and explanation of methodology.

STUDY	METHODOLOGY	FINDINGS	LIMITATIONS
	Sample – n=85 intervention, n=81 control group. **Soundfield type**- unknown **Acoustic measurements** – not recorded **Hearing Screening** – not performed **Measures** - The Informal Reading Inventory (IRI) was used to measure literacy achievement at the middle and end of the year. Questionnaire -Teacher Appraisal of Listening Difficulty	The increase in scores was greater for young people that had English as a second language or had identified additional support needs. Teachers provided strong support for the positive change made by the soundfield system.	Not peer reviewed. No data on selection of the control and intervention group No hearing screening performed No comparisons made between the control and intervention groups Baseline assessments not completed prior to intervention Teachers not blind to the intervention.

STUDY	METHODOLOGY	FINDINGS	LIMITATIONS
MASSIE AND DILLON (2006A) **MASSIE AND DILLON (2006B)**	**Design** - Within subject crossover. 8 of the classes (1-8) were exposed at different times to amplified, unamplified, a single microphone and two microphone conditions. Four classes (9-12) were exposed to single and dual channels for two terms. **Sample**- n=242 (114=female, 128=male) Mean age=6 years, 8 months. **Soundfield** - A NAL Twin FM soundfield amplification system **Acoustic measurements** – ambient noise and RT_{60}	Improvements in literacy and numeracy were observed. Teachers observed significant improvement in attention, communication and classroom behaviour when using the soundfield. Students responded to improved listening experience.	Measures used to determine the educational outcomes for the students were non-standardised. Teachers involved were not blind to the outcomes and so there was the potential for response bias.

STUDY	METHODOLOGY	FINDINGS	LIMITATIONS
	Hearing Screening – Pure tone audiometry (PTA) **Measures** – reading, literacy and numeracy. The Queensland's early identification and monitoring of students that are having difficulties with numeracy and literacy. Measures taken at the start, middle and end of year. Questionnaire – Teacher Opinion re Performance in classroom (TOPIC). Whole class open-ended questions for the students.		
HEENEY (2007)	**Study Design -** Quasi-experimental longitudinal study.	Significant improvement was recorded in the intervention classroom: listening and reading comprehension,	No hearing screening performed.

STUDY	METHODOLOGY	FINDINGS	LIMITATIONS
	Sample – n=626 (n=298 female, n=328 male), n=436 intervention, 187=control. **Soundfield** - Phonic Ear Easy Listener four speaker soundfield with a boom microphone. **Acoustic measurements** – not recorded. **Hearing Screening** – not performed. **Measures** - National assessments of listening comprehension, reading comprehension, reading vocabulary, phonological skills and mathematics. Teacher and student questionnaire.	reading vocabulary and mathematics. Significant improvement in phonological skills assessments for students aged 5-6 years old. Positive teacher response. Quieter classrooms were observed, improved behaviour and understanding of instructions. Less vocal strain. Students found it easier to hear and had improved the harmony of the classroom.	No acoustic measurements recorded to determine differences between conditions. Intervention group randomly assigned but control group selected on availability. No comparisons made between the control and intervention groups. No sample size calculations. No detail on attrition rates. Only certain students completed questionnaire –

STUDY	METHODOLOGY	FINDINGS	LIMITATIONS
			no information on the handing of missing data.
Millett and Purcell (2010)	**Study Design -** Quasi-experimental, non-equivalent groups longitudinal study. **Sample** – n=486 (n=231 female, n=255 male), n=247 intervention, 239=control. **Soundfield** - Phonic Ear VocaLight infrared **Acoustic measurements** – not recorded. **Hearing Screening** – Only 321 students screened using otoacoustic emissions. **Measures** - Developmental Reading Assessment presented and scored by	The results indicated greater changes in the total percentage of students with improved reading grades at the end of the school year in rooms using soundfield. Improved scores for students identified as at risk for reading difficulties. Both results were not statistically significant. Positive teacher response to the use of soundfields. Less repetition, improved	Not all participants received hearing assessments. No data on what constituted a pass. No data on selection of the control and intervention group. No comparisons made between the control and intervention groups. Teachers not blind to the intervention and they marked and presented the

STUDY	METHODOLOGY	FINDINGS	LIMITATIONS
	each student's classroom teacher. Teacher questionnaire - Opinion and Observation List.	classroom management and less vocal strain.	assessments. Also, they completed satisfaction questionnaires. No acoustic measurements were taken. Initially 514 students included in the study, no data on attrition and handling of missing data.
WILSON ET AL. (2011B)	**Study Design –** Repeat measure longitudinal study. **Sample** – students n=147 (n=70 female, n=77 male), teachers n=8 (n=8 female) Student age – Mean= 8 years, 2 months)	Significant improvement in student listening ($p<0.01$) and auditory analysis ($p<0.05$) auditory analysis in the brick building with a solid wall separating them from neighbouring classrooms.	No hearing screening performed and so there could be confounding variables. No data on the selection of the participants and

STUDY	METHODOLOGY	FINDINGS	LIMITATIONS
	Soundfield - single-speaker Redcat devices with a two Lightmic microphones **Acoustic measurements** – Unoccupied background noise levels and RT_{60}. **Hearing Screening** – No performed. **Measures** – Three different classroom environments: brick building with neighbouring rooms separated by solid walls; brick building separated by open space and demountable buildings, separated from other classrooms by a solid wall. The effectiveness of soundfield in these three different environments measured using The Literacy and Listening Index, Test of Auditory Analysis	This has the lowest background noise levels and second lowest RT_{60}.	allocation to the different conditions. No data on the use of the additional microphone and if this was evenly used in all conditions. No sample size calculations.

STUDY	METHODOLOGY	FINDINGS	LIMITATIONS
	and Listening Inventory for Education: Student Appraisal of Listening Difficulty.		
DOCKRELL AND SHIELD (2012)	**Study Design –** Prospective cohort – 6 months follow up. **Sample** – n=740 completed baseline questionnaires. n=478 completed follow up questionnaires. N=393 completed both. **Soundfield** – not recorded **Acoustic measurements** – RT obtained in some classrooms. Questionnaires **Hearing Screening** – not performed. **Measures** - Modified versions of standardised tests for spelling,	No significant difference between the students in soundfield and non-soundfield rooms in areas such as numeracy, reading, or spelling. There were larger gains in performance for students in soundfield rooms for speed of language processing and listening comprehension. Improved scores in rooms with poorer acoustics. No	No hearing assessments performed. High attrition rate reduced sample size and so potential bias. Also, no discussion on the handling of missing data. Allocation to the control and experimental groups unclear. Poor training level of staff involved.

STUDY	METHODOLOGY	FINDINGS	LIMITATIONS
	numeracy, and speed and accuracy of nonverbal processing, the Suffolk Reading Scale and listening comprehension test were deployed to measure academic and cognitive ability. Teacher and student questionnaires.	significant benefit in schools with better acoustics.	
VICKERS ET AL. (2013)	**Study Design –** Within subject repeat measure. **Sample** – n=44 (n=20 female, n=24 male). Year 2 (age 6 years, 11 months- 7 years, 10 months). Year 3 (7 years, 11 months-8 years, 10 months). **Soundfield** – Digimaster 5000 Dynamic soundfield.	Vocabulary age had an effect on the young person's ability to hear the speech signal in noise. The soundfield lead to an improvement in speech perception. The biggest impact was for young people with lower vocabulary ages.	No hearing screening performed only an examination of the ears and test of middle ear compliance. Even young people that received negative responses to tympanometry remained in the sample.

STUDY	METHODOLOGY	FINDINGS	LIMITATIONS
	Acoustic measurements – RT_{60} recorded. No noise surveys completed. **Hearing Screening** – Otoscopy and tympanometry. **Measures** – Modified and shortened version of the CHEAR auditory perception test. Run twice separated by 2 weeks.		No SNR ratios achieved at the individual student level. RT_{60} recorded but no information on C_{50} which could influence the outcome of speech assessments.

Table 1: Main findings and limitations of previous studies into soundfield systems in education.

2.5.2 Academic measures and educational outcomes

Arnold and Canning (1999) used a controlled clinical study to measure the effects of a Frequency Modulation (FM) soundfield system on reading comprehension in mainstream primary classrooms. Forty-nine young people (M=9.92 years) from two upper un-streamed classes completed a modified version of the Neale Analysis of Reading Ability (NARA) assessment in an amplified and unamplified environment. The RT=2.04s was calculated using Sabine's equation. Noise measures were recorded for unoccupied (M=60.05dBA), occupied (M=66.67 dBA) and soundfield system in use (M=73.17 dBA). The NARA was administered using an audio recording with the students required to answer in writing. Three levels of the test were administered: Level 1, Level 2 and Level 3. The results indicated a significant improvement in comprehension over the three levels of the NARA in the rooms provided with soundfield. The differential between the amplified and non-amplified was greater as the complexity of the NARA increased. There was no gender difference in the score results. In the pupil questionnaire, 54 per cent recognised an improvement in their listening due to the soundfield system whilst 44 per cent felt that there was no change. There were several shortcomings in this study. There was no hearing screening before the intervention. Furthermore, the participants were required to write their response to the NARA assessment and so this would disadvantage learners that had literacy issues. No pre-intervention assessment of language levels was undertaken. It is unclear if those marking the assessments were blind to the intervention.

Darai (2000) also compared literacy levels using a pre and post repeat measure methodology. The dependent variable was the Informal Reading Inventory and the *Teacher Appraisal of Listening Difficulty* inventory, which is a subsection of LIFE. This was used to identify changes in attention, classroom participation, and learning. Four schools participated with each assigned an intervention and control classroom. Eighty-five young people were part of the intervention group and eighty-one were in the control group. The results suggest that the intervention classrooms achieved greater literacy gains than the control. The improvement in scores was greater for young people that had English as a second language or had an identified additional support need. The LIFE inventory provided strong support for the positive change made by the soundfield system. This paper was unpublished and so was not subject to peer review. There is limited data on the methodology about recruitment, sample size, method of allocation to the intervention and control conditions. Furthermore, there is no data on the numbers of students with identified additional needs or any baseline assessments performed. There is no information on the statistical methods used to analyse the data. No hearing screening, measurement of background noise levels or RT was performed.

Millett and Purcell (2010) examined changes in reading outcomes for 486 grade 1 students (247 in the intervention and 239 in the control classrooms) in twenty-four classrooms (12 intervention and 12 control) over a one-year period. At the pre-intervention stage, 321 students had their hearing screened using otoacoustic emissions with 43 receiving refer outcomes. The Developmental Reading Assessment was the pre and post-test assessment which was presented and scored by the classroom teacher. In addition, the *Teacher Opinion and Observation List, Voice* subsection of LIFE was used to assess the impact of the soundfield system. The results indicated a positive difference in reading grades in the intervention classroom, although the results were not statistically significant. Students identified at being at risk for reading difficulties showed the biggest improvement. Teachers were very positive towards the soundfield system. However, not all the students received a hearing screen due to time and financial constraints and of those screened the refer outcome was high at 13.4%. Otoacoustic emissions are generally used to screen babies with the global refer rate between 0.5-5 per 1000 cases and so these results indicate a high false-positive rate. This could mean participants were incorrectly excluded from the sample which could introduce bias (Dada et al., 2013). Five hundred and fourteen participants were initially included with 28 students moving school however there is limited information on the method used to handle missing data. The teachers graded the papers and commented on the effects of the soundfield, but they were not blind to the intervention.

Massie and Dillon (2006a) also examined the effectiveness of soundfield on three educational outcomes: reading, literacy, and numeracy using Queensland's *Early Monitoring and Assessment of Numeracy and Literacy*. The participants were twelve classes of 242 Year 2 students (M=6 years, 8 months). The majority of students were from English as an additional language background. A within-subject crossover design was deployed with each class acting as the control. The results indicate that there was an improvement in reading, writing and numeracy scores irrespective of whether English was an additional language. The main shortcomings of this study were the non-standardised methods used to measure educational outcomes. Furthermore, the class teachers graded the papers but were not blind to the intervention with the potential for response and expectation bias.

Wilson et al. (2011b) assessed the effectiveness of soundfield on listening, spelling, reading comprehension and auditory analysis in three different classroom environments. A pre and post repeat experimental measure was used with control and intervention classrooms in three different types of school building: brick separated by solid walls, brick with an open space separation and demountable buildings separated by solid walls. The control and intervention class in each school were selected randomly. The T_{mf} in two classes was <1s and the other two >1s. The background noise levels in all classes were above the recommended guidelines. The results demonstrated a small but statistically significant improvement in listening and auditory analysis but only in the brick-built schools separated by solid walls. This was the classroom with the lowest level of background noise (47-50 dB (A)) and second-lowest RT_{60} (0.87-0.91s).

Research by Dockrell and Shield (2012) involved 22 classrooms in eight schools (14 were intervention and eight control). A prospective cohort study design was used with a six-month follow-up. The study design included acoustic surveys of classrooms, noise questionnaires completed by students and teachers and modified versions of standardised tests for spelling, numeracy, reading, listening comprehension and speed and accuracy of nonverbal processing. Teachers with a soundfield installed also completed a questionnaire on usage as the system did not incorporate Datalogging. The results indicated that there were no significant improvements in numeracy, reading, or spelling for rooms fitted with a soundfield. The increased performance was recorded for the speed of processing and listening comprehension. Of the five schools that had soundfield installed, three had shorter T_{mf} ($\leq$0.52s) and so would be defined as having good acoustics for speech. The analysis showed that the soundfield was more effective in poorer acoustic conditions. This is consistent with Wilson et al. (2011b) where improved outcomes were observed in amplified classrooms with longer RT (0.87-0.91s). Dockrell and Shield (2012) also found there was no significant difference in the awareness of noise or listening experience in both the control and intervention conditions. One of the limitations of this research was the high level of attrition as 740 students completed the baseline questionnaire, 478 completed the follow-up and only 393 completed both baseline and follow-up. In addition, 30 percent of the teachers had stopped using the soundfield at the six-month follow-up. Attrition can reduce power, introduce bias and compromise validity (Amico, 2009). Furthermore, there was no discussion on the methods used to handle the missing data. Not all the classrooms received an acoustic survey and no hearing assessments were performed before the intervention.

The quasi-experimental longitudinal study by Heeney (2007) is the only study to factor social deprivation into the research design. Academic performance was measured using standardised national assessments of listening comprehension, reading comprehension, reading vocabulary, phonological skills and mathematics. A teacher questionnaire was also completed. The state schools in New Zealand are rated on a socioeconomic scale from one to ten; one represents the most disadvantaged and 10 the least disadvantaged. One school was selected from the following socioeconomic scales: 1, 2, 5, 6 and 10. Six different year groups were involved. The results from the longitudinal study found that there was an improvement in listening and reading comprehension in the groups exposed to soundfield technology and the same was true for phonological skills assessments for students aged 5-6 years old. Year groups 5 and 6 also showed an improvement in standardized mathematical scores (there were no standardized assessments in mathematics for the other year groups). Differences in listening comprehension scores were recorded between schools in the lower and higher socioeconomic groups but it was not statistically significant. There was an overall positive response from the teachers to the soundfield equipment.

Although a large sample was followed in the study (see Table 1), it covered a broad stage of the schooling system and there was no data provided on the statistical significance of the sample for each year group in the control and intervention classrooms. There was only one school from each decile rating and so the significance of the sample size for each socioeconomic group was not explained. This is particularly relevant as there is a large difference between the number of students in the control and intervention classes, 30 percent and 70 percent of the sample respectively. Also, the decile rating was for each school and was not specific to each student and so the socioeconomic status of each participant was unknown. There appears to have been no attrition on this yearlong study which appears unusual. Qualitative data was collected through the teacher questionnaires, but they were not blind to the intervention and consequently, this was susceptible to response and expectation bias. There was no hearing screening performed before the start of the longitudinal study and no measurements of ambient noise or reverberation were taken. No comparisons of the standardised assessment results were made between the intervention and control groups, therefore, other extraneous factors may have influenced the outcomes

2.5.3 Questionnaires and educational outcomes

McSporran et al. (1997) deployed a before and after methodology to assess the listening and attending behaviour of young people in two mainstream primary classrooms. The selection of the participants was made by the class teachers who completed the Screening Instrument for Targeting Educational Risk (SIFTER) which consists of fifteen statements covering academic performance, attention, communication, class participation, and school behaviour. The effectiveness of the intervention was measured using the Children's Auditory Processing Performance Scale (CHAPPS) which was also completed by the class teacher. Six different categories were measured: noise; quiet; multiple inputs; auditory memory, sequencing; auditory attention span. The results from the study indicated there was a significant improvement in listening behaviour according to their CHAPPS scores. Participants identified with the greatest difficulty in the SIFTER showed the greatest improvement in their CHAPPS score. As in previous studies, the teachers rated performance but were not blind to the intervention. All participants were exposed to the soundfield with no control. There was also a lack of data regarding how often the teachers involved in the study used the soundfield system.

A longitudinal study carried out in two phases was completed by Rosenberg et al. (1999). The sample size and equipment installed are presented in Table 1. Ambient noise measurements were taken from five positions in the room but there was no RT. Teachers in both phases completed an *ICA Classroom Description Worksheet* on noise sources, classroom treatments, and classroom design. The *Listening and Learning Observation* questionnaire was used to measure listening behaviours and academic outcomes. In addition, a randomly selected ten students in each class were also assessed using the *Evaluation of Classroom Listening Behaviours*. Paired *t*-test results showed a significant difference in scores between the experimental and control group. The greatest improvement was in the younger kindergarten age group and the poorest scores recorded in the first and second-grade students in the control group. Response and expectation bias cannot be discounted as the teachers graded the questionnaires. All had received training prior to the study on the benefits of the soundfield system. The confounding variables include the teachers determining the type of microphone used with the system, only a partial number of students had their hearing screened in phase 1, no student was tested in phase 2 and the RT in the rooms was not recorded.

2.5.4 Dynamic soundfield and speech recognition

Only one study to date has tested the effectiveness of dynamic soundfield on speech perception within a classroom. Vickers et al. (2013) used a modified and shortened version of the closed-set CHEAR Auditory Perception Test (CAPT) speech assessment. The school was located in an urban area where the majority of learners were not from families with English as the home language and as such the vocabulary age was tested prior to the intervention. The CAPT test was presented on an interactive whiteboard and the students responded using a personal response hand-held voting system. The speech sound was calibrated at 1 metre from the signal and this was recorded at 65dB(A) and the noise level measured at the centre of the room was 46dB(A). The older year group had an additional noise condition set at 52 dB(A). The speech test was presented randomly with and without the dynamic soundfield. Overall the results suggested an improvement in performance for all students with dynamic soundfield but that the results were stronger for students with lower vocabulary levels. One of the limiting aspects of this study is the small sample size. Furthermore, the students did not have sufficient time to become accustomed to the dynamic soundfield before the testing. In addition, no hearing screening of threshold levels was performed.

2.6 Overview and limitations of previous studies

The overall aim of this review was to set out the conceptual model underlying the experience that young people have of listening in noise and to place this in the context of a contemporary classroom. Predictive factors for the detrimental effects of listening in noise are age, experience, language and vocabulary levels (Nittrouer and Burton, 2005, Vickers et al., 2013). The review of the eleven previous studies has reinforced the need for research into the effects of dynamic soundfield technology in delivering improvements in educational attainment and reducing the attainment gap between learners living in the most and least deprived areas of the country. In general, previous studies have been compromised by limitations in the study design. When designing and formulating the aims for this research, which are set out in Chapter 3 the following factors are considered:

- There is a paucity of research into the effects of dynamic soundfield on attainment that use standardised measures of assessments that are implemented nationally in schools.
- There is also no published research to date that measured the effectiveness of dynamic soundfield using standardised assessments that were marked by an external body, blind to the intervention.
- Previous studies that factored the daily use of the soundfield system into the study design have done this through questionnaires. These were completed by the teacher who was not blind to the intervention and so was susceptible to response and expectation bias. There has been no study that has used Datalogging to objectively record the amount of time each day the soundfield system was in use.
- Previous study designs have not managed confounding variables by including a comprehensive range of audiological and acoustic measurements: hearing, ambient noise, occupied noise levels, RT_{60} and speech clarity (C_{50}).

- The questionnaires used in previous studies about the teachers' and young persons' perceptions of soundfield amplification was susceptible to response bias. The training provided in some of the studies also discussed the benefits of the system and so introduced bias.
- There is a paucity of longitudinal research into the effects that soundfield has on learners living in areas of social deprivation. Only one study factored this into the study design, however, there is limited information on the methods used to calculate the sample size. Furthermore, the deprivation status was established at a school rather than an individual level (Heeney, 2007).

The research aims and methods are discussed in the next chapter.

Chapter 3 General Methodology

3.1 Introduction

The general methodology regarding ethics, recruitment of participants and schools, equipment selection, pre-intervention assessments, and research instruments will be discussed in this chapter. Figure 8 presents the main components of the study design.

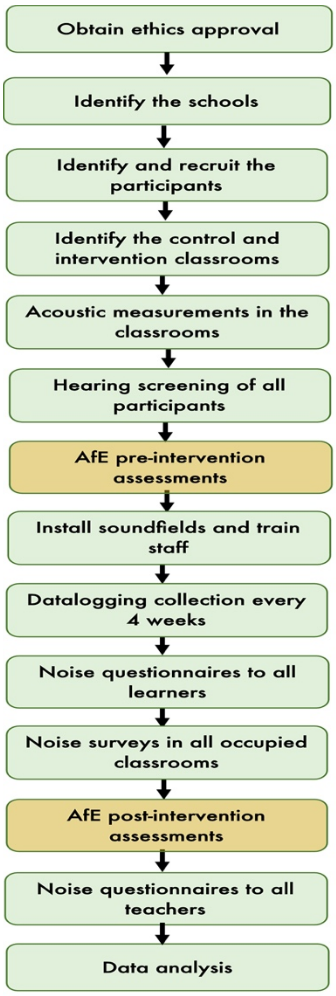


Figure 8: Flow chart showing the major components of the study design.

3.2 Research aims

In broad terms, the general study design set out in this chapter seeks to address the paucity of research in this field and address the limitations of previous studies.

3.2.1 Primary aim

1) To determine if there were significant improvements in educational attainment for Primary 3 learners exposed to dynamic soundfield amplification in their mainstream classrooms.
2) To establish if the gap in educational attainment between Primary 3 learners from the most and least deprived quintiles reduces after being exposed to dynamic soundfield.
3) To investigate if the changes in scores for each of the subtests and modules were significantly different between SIMD 1 learners exposed to dynamic soundfield and SIMD 1 learners in the control.
4) To investigate if the changes in scores for each of the subtests and modules were significantly different between SIMD 5 learners exposed to dynamic soundfield and SIMD 5 learners in the control.
5) To investigate if the efficacy of dynamic soundfield is moderated by the speech weighted C_{50} properties of the classrooms.

3.2.2 Secondary aim

1) To compare and characterise the acoustic environments in the control and intervention classrooms.
2) To compare the characteristics of the participants to the general population to establish internal and external validity.

3) To establish quantitative data on the overall use of the dynamic soundfield system in each intervention classroom.
4) To explore the views of learners and teachers on the use of the dynamic soundfield in the classroom.

3.3 Ethics

As this research involved primary 3 learners in Fife schools, ethical approval was required from both Fife Council and the University of Edinburgh. Any research in Fife schools requires the consent of the Research and Ethics Committee and this was granted after the submission of the research proposal and a meeting with the committee. See Appendix 1 for a copy of the approval letter.

All primary 3 learners sit the AfE (InCAS) assessment at the end of the summer term. The additional commitments required by the young learner as a result of this research were:

1. A hearing screening assessment undertaken in school prior to the intervention.
2. A pre-intervention AfE (InCAS) assessment at the start of term.
3. The completion of a noise survey questionnaire.

In addition to the above, the teachers and schools committed to the following:

1. Whole class noise surveys.
2. Acoustic measurements of the control and intervention classrooms.
3. Datalogging collected from the intervention classrooms a maximum of every 30 days.
4. Copies of weekly timetables.

5. Installation of a dynamic soundfield system (intervention classrooms only).

As this was regarded as research that covers novel procedures and topics of a more sensitive nature a Level 2 ethics application was made to the University of Edinburgh. The following measures were included in the application to protect privacy, ensure that no participant was disadvantaged and obtain the informed consent of the participants:

1. The parents/carers of the learners were provided with a letter about the research, asked to discuss it with their child and then given the option to opt-out.
2. The learners in the control and intervention classrooms were provided with an information leaflet in an accessible format which was discussed with the whole class. The learners signed the leaflet to give informed consent. Additional visits were arranged to obtain the consent of learners absent at the time of the initial discussion.
3. Headteachers and class teachers in both the control and intervention classes were given an opt-in letter which was signed.
4. The concealed random allocation of the control and intervention classrooms resulted in some learners being placed in the control group. Dynamic soundfield was considered an intervention that could deliver improvements in educational outcomes. To ensure no participants were disadvantaged learners in the control group were provided with a dynamic soundfield system the year following the intervention.
5. All electronic data was kept on an encrypted laptop that was password protected.
6. All the consent forms were retained in a locked file.
7. The schools were allocated a number and were not to be identified in the final thesis.

8. No individual was to be identified in the final thesis.

All participants were advised that they could withdraw at any time from the research without giving reasons. Approval was granted, and copies of the consent forms are provided in Appendix 2.

3.4 Participant Selection Criteria

Participants were identified using six inclusion criteria:

1. The learners attend mainstream primary 3 classes in the Fife region of Scotland.
2. The learners had typical hearing threshold levels as defined by the British Society of Audiology (BSA Professional Practice Committee, 2015)
3. The participants attend school buildings with similar construction types and with enclosed classrooms.
4. The internal configuration of the school ensured that the control and intervention classrooms were not adjacent to dinner halls, gyms or assembly halls.
5. The school roll was ≥170.
6. The sample size for statistical significance consisted of learners from the 20 per cent most and least deprived areas of Scotland as defined by the Scottish Index of Multiple Deprivation (SIMD) (Scottish Government, 2014b).

The Scottish Index of Multiple Deprivation is the Scottish Government's official method for measuring areas of multiple deprivation (Scottish Government, 2014b).

There has been criticism of the SIMD as a method of allocating state funding to meet rural poverty and health inequality (McKendrick et al., 2011, Fischbacher, 2017). Furthermore, Paterson et al. (2019) regarded it as a crude tool to measure widening access to higher education. However, the SIMD has achieved recognition by the Royal Statistical Society and is used nationally by government organisations and is incorporated into the pupil census. Furthermore, for the purposes of external validity, it allows a comparison between the participants and the target population.

3.4.1 Rationale for participant selection criteria

The primary 3 age range were chosen as they generally listen less selectively than adults, have poorer attentional skills and are less familiar with language structures which makes them more susceptible to informational masking (Eisenberg et al., 2000, Stollman et al., 2004, Blandy and Lutman, 2005, Wightman et al., 2010). Young people who did not have typical hearing threshold levels were excluded from this study as both conductive and sensorineural forms of deafness are associated with reduced ability to perceive speech in noise. Excluding deaf learners would control for potential confounding variables and so minimise alternative explanations for the outcome measures observed.

Schools with open-plan classrooms were excluded from selection as soundfield may not be as effective in such settings. Open-plan classrooms are more susceptible to intrusive internal noise from other classes in the vicinity and the larger teaching spaces may have longer RT_{60} (Wilson et al., 2011a). Similar construction types were an inclusion factor as this along with room volume, shape and classroom finishes are significant predictors of both the RT_{60} and C_{50} values. Control and intervention classrooms in the schools were not adjacent to gyms, lunchrooms or assembly halls to minimise the levels of intermittent noise internal to the school but external to the classroom. Schools with a roll of <170 were excluded as there was a higher probability of multiple composite classes in smaller schools. Furthermore, selecting larger primary schools made it feasible to have more of the control and intervention classes within the same school.

Primary 3 age learners from the 20% most (SIMD 1) and least (SIMD 5) deprived areas of Scotland were selected as there is a poverty associated educational attainment gap between these two quintiles (Scottish Government, 2018a). As revealed in Chapter 1, there is a gap in the curricular areas of reading, literacy, listening and talking and numeracy. Furthermore, the longitudinal research study, *Growing Up in Scotland* funded by the Scottish Government examined the cognitive ability of young people between the age of 34 months and 58 months. It found there was a significant gap in problem-solving ability and knowledge of vocabulary between young people in the lowest income group compared to the highest income group. Age equivalent scores indicated there was a 13-month gap in vocabulary knowledge and a 10-month gap for problem-solving (Bradshaw, 2011). This is consistent with the research of Nittrouer and Burton (2005) who observed that young people living in deprivation have delayed phonemic structure development which impairs the process of storing and retrieving words from working memory. Young people from SIMD 1 quintile were selected for this research as it can be hypothesised that they will be more affected by noise as they do not have the language knowledge to resolve ambiguities caused by masking and so may benefit from dynamic soundfield amplification.

3.4.2 Sample Size

The methodology used to calculate the sample size required to test H_0 and answer specific research questions is an important factor in any experimental design due to its effect on statistical power. Haas (2012) defines statistical power as the probability of detecting the difference between the control and intervention groups when that difference actually exists. A sample size calculation should identify the number of participants required to correctly reject or accept the H_0 and so avoid Type I and Type II errors. Type I errors are associated with false-positive outcomes in which the H_0 is incorrectly rejected. To reduce the odds of a Type I error occurring a statistical significance level of $\alpha=0.05$ is conventionally applied as this determines there is a 95% certainty that the outcome observed was not by chance (Prajapati et al., 2010). Type II errors are associated with false-negative conclusions in which the H_0 is incorrectly accepted. A $1-\beta=0.20$ is conventionally applied which means there is a <20% likelihood of a false negative outcome (Noordzij et al., 2010).

Another key element of sample size calculation is the standard deviation of the population being studied. The AfE (InCAS) assessments provide standardised scores with a mean of 100 and a standard deviation of 15 (Centre for Evaluating and Monitoring, 2014). The fourth element of a sample calculation is the size of the effect that is considered relevant. For quantitative data, the effect size (δ) represents the mean difference between the control and intervention groups or the difference in mean scores at the pre-test and post-test stage (μ_1-μ_2), divided by the grouped standard deviation (σ). Suresh and Chandrashekara (2012) recommend using previous studies to calculate the effect size and if this information is not available then it should be determined from a review of the literature and logical assertion.

As discussed in Chapter 2, previous studies into soundfield have a dearth of information on sample size calculation and effect size. Meta-analysis is a quantitative epidemiological study design for systematically assessing the findings from previous research to derive conclusions about the field of study (Haidich, 2010). Hattie et al. (1996) examined innovations in education and academic achievement across 304 meta-analyses, based on 40,567 studies. A detailed study of the results illustrates that the typical effect size for interventions involving the classroom environment had an overall $\delta=0.56$. This is relevant for this research as dynamic soundfield is an intervention intended to improve the listening environment. Adopting a more conservative effect size of $\delta=0.5$ would ensure that the same size is not underestimated. In AfE (InCAS) this equates to a change in the mean of 7.5. Using the following formula, a sample size of 63 is required from each of the SIMD 1 and 5 quintiles in the control and intervention groups based on a repeat study design (Allen, 2011).

$$n = \frac{2\,(Z_{\alpha/2} + Z_{1-\beta})^2}{\left(\frac{\mu_1 - \mu_2}{\sigma}\right)^2}\ :$$

$$n = \frac{2(1.96 + 0.84)^2}{\left(\frac{7.5}{15}\right)^2} = \frac{15.68}{0.5^2} = 63$$

To account for possible attrition, a 10 per cent addition is advised to any sample size calculation, providing a target sample of 69 learners from SIMD 1 and 5 in each condition (Israel, 1992).

3.4.3 Deprivation status of participants

The method used to establish the deprivation status of the participants was the Scottish Government's Index of Multiple Deprivation. The participants' postcodes were obtained from SEEMiS, which is the national education data information system for Scottish schools (SEEMiS, 2017). These were inserted into the SIMD Postcode Checker. This Excel file converts postcodes into datazones, along with their associated ranks, quintiles, deciles, and geographies (Scottish Government, 2014b). Quintile 1 is the most deprived area and 5 the least deprived. The individual learners were categorised based on the quintile rating of their home postcode. The quintile rather than decile was used in this research as it made it possible to have a manageable sample size.

3.5 School selection

Fife Council provided a list of the 141 primary schools in the region, which were ranked on a measure of the SIMD; based upon the home postcodes of the learners attending the school. SIMD divides Scotland into 6,505 areas called datazones, with each representing around 350 households. Seven domains are used to establish the deprivation status of the datazones: income, health, access to services, employment, education, housing and crime (Scottish Government, 2014b). The 2012 index formed the basis of this research and incorporated minor amendments to the indictors to reflect welfare reform and changes to the female state pension age (Scottish Government, 2014b).

Each school in Scotland is ranked based on the datazones with 1 signifying the most deprived school and 6,505 the least deprived. The school with the highest datazone ranking in Fife was 5936 and the lowest was 804. As Figure 9 illustrates, initially 70 schools, 35 from the least deprived (datazone ranking: 4343 to 5936) and most deprived areas (datazone ranking: 2148 to 804) were identified as having a catchment that would include a significant proportion of learners from SIMD quintile 1 and 5. Fife Council applied a capped learner to teacher ratio of 1:18 for primary 1-3 classes in areas of social deprivation and so to achieve an appropriate sample size more schools in areas of social deprivation were included. To achieve a sufficient sample size, it was anticipated that between 10 and 14 schools would be required

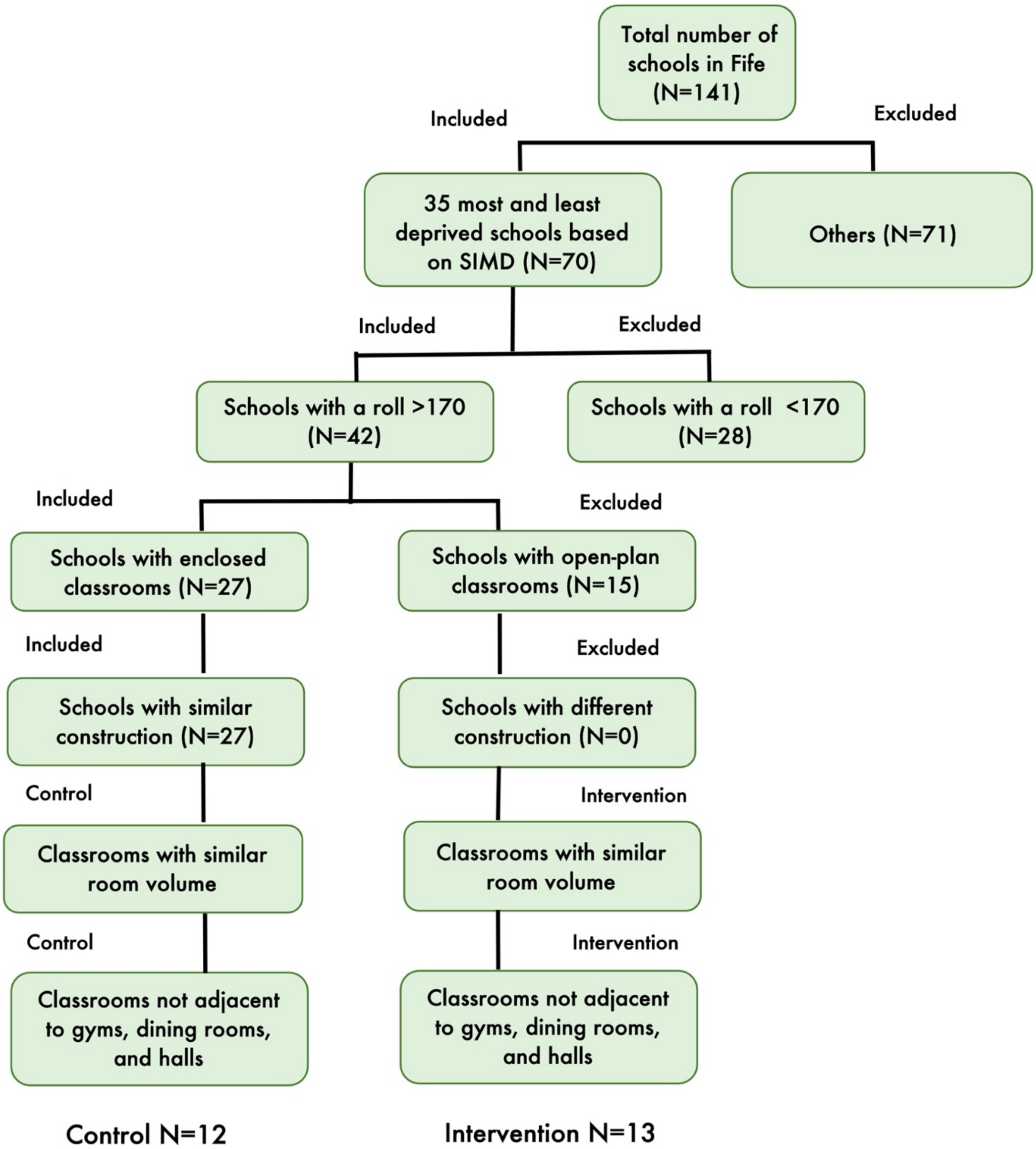


Figure 9: Methods used to identify the schools that met the inclusion criteria.

The school roll for each of the 70 schools was checked against the Scottish Pupil Census of 2012 and 28 schools were excluded as they had rolls <170: twenty schools were from the least deprived areas and eight from the most deprived

(Scottish Government, 2012b). Fife's Property Asset Management Strategy Team provided information on the classroom configuration and school construction types for the 42 primary schools identified as having a school roll of ≥170. At this stage, 15 schools were excluded from selection as they had open plan classrooms – eight from the least deprived and seven from the most socially deprived areas.

There are three common building types in the larger primary schools in Fife:

1. Those constructed from a reinforced concrete frame, concrete floors and roof brickwork internals.
2. Those built in the 1920s from traditional brick walls, timber floors, timber slated roof, and brickwork internals.
3. Victorian buildings constructed of stone.

The net area of the classrooms in the schools ranged from 65.37 to 80.44m^2. Internal plans of the schools were provided to ensure that the self-contained classrooms were not adjacent to gyms, lunchrooms or assembly halls. Figure 9 illustrates the number of schools that met the inclusion criteria was 27: 14 in the least deprived and 13 in the most deprived areas of Fife. The plans of the schools that participated in the research are provided in Appendix 3.

Meetings with the Education Officer and statistician for Fife Council were arranged and following this it was agreed an email would be sent to the headteachers of the schools that met the inclusion criteria. An email was sent in the spring term and a follow-up telephone conversation with all headteachers took place over the next five days. After a request for more information, a meeting with eight of the schools was arranged. Of the 14 schools in the least deprived areas, one school could not participate as the headteacher was retiring and a new one was yet to be appointed, one withdrew without a reason after initially agreeing to participate and five schools did not respond to the email or telephone conversation. In addition, one of the schools had historically been associated with the Royal Air Force (RAF) and due to a change in the base from RAF to the army; it was not thought appropriate for it to take part, as the school roll would fluctuate throughout the year. Six of the schools in the least deprived areas agreed to participate. Of the 13 schools in the most deprived areas, two schools could not participate as their headteacher was retiring in the summer and no new appointment was made, two schools were earmarked for closure and one school's roll had dropped and there was uncertainty about teaching numbers. One school did not respond. Seven schools agreed to participate. Overall, thirteen headteachers from thirteen schools agreed to participate in the research.

3.5.1 Control and intervention classroom selection

As Figure 10 illustrates a longitudinal repeat measure design was used in this study. At the end of the summer term, the Senior Management Team in the research schools allocated the learners and teachers to each class blind to which was the control and intervention conditions (see Chapter 4 for a fuller discussion on this process). Once the classes and teachers had been allocated, a survey of the classroom took place with the researcher blind to which learners and teaching staff had been assigned by the school. The rooms fitted with a dynamic soundfield became the intervention classrooms and those without any amplification system became the control. The reason for the comparison between the control and intervention groups is that without a comparison it is not possible to determine if the outcome is the consequence of the intervention or the product of a confounding variable (Dockrell and Shield, 2012).

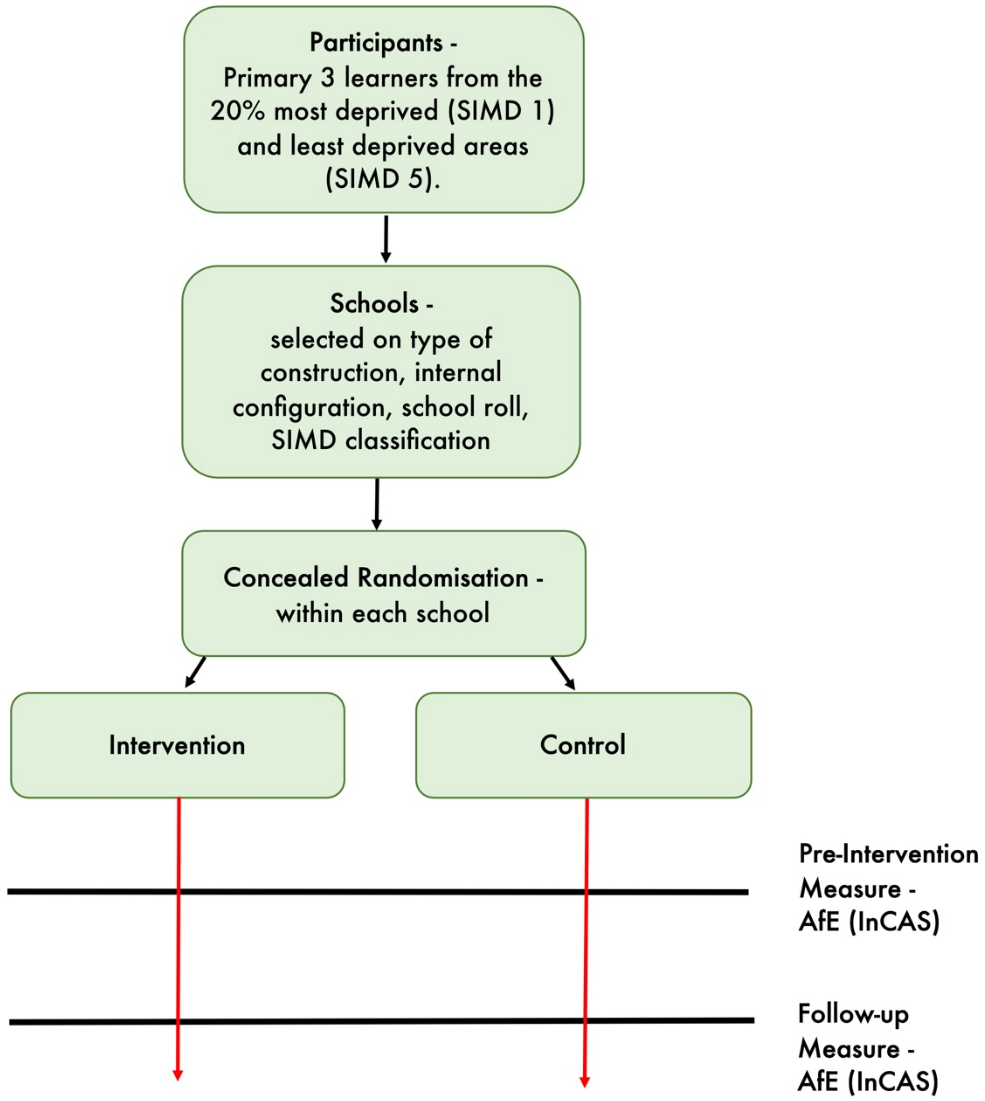


Figure 10: Repeat design method and the concealed random allocation of the control and intervention classrooms used in this research.

The inclusion criteria for the intervention classrooms:

1. A room that could accommodate a bracket at the front of the room, 1.4 meters from the ground as stipulated in the fitting guidelines.
2. A room that did not have asbestos in the wall as drilling would be prohibited.

3. A power socket was close to the front corner of the room to provided electric power for the soundfield.

Where more than one class met the inclusion criteria, a random allocation determined the intervention room. In schools that only had a single primary 3 class by default became the intervention classroom. Due to previous studies indicating that there was a high attrition rate amongst teachers using the soundfield system a greater number of intervention classes than control classes were incorporated into the study design (Dockrell and Shield, 2012). Each of the 13 schools had an intervention class, two schools did not have a control as there were not a sufficient number of primary 3 students to form two classes. Two of the research schools had two control classrooms. One reason for a pre-test/post-test design involving 25 control and intervention groups was that it would reduce the threat to internal validity created by maturation as both groups would be subject to the same process.

3.6 Intervention Equipment

3.6.1 Soundfield Amplification Equipment Selection criteria

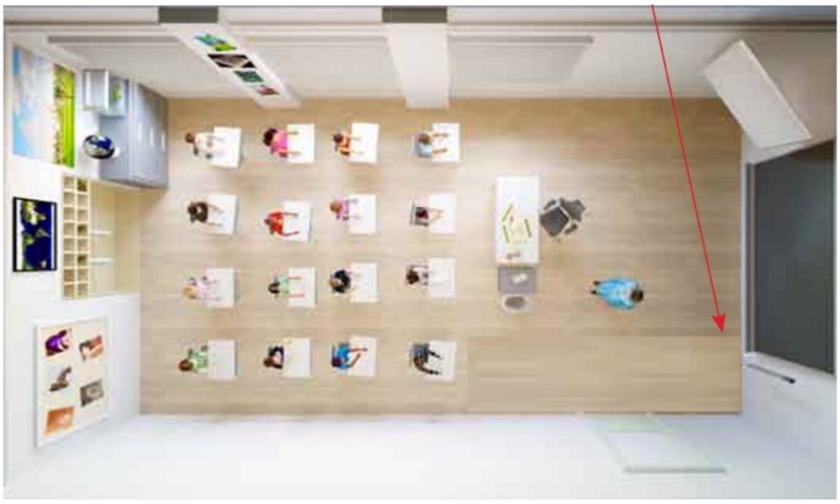
The soundfield was selected using five selection criteria:

1. The soundfield system could collect and store data on usage (Datalogging).
2. The system should provide a variable level of gain to the voice of the teacher/learner.
3. A single speaker location rather than the requirement for cabling to multiple speaker locations.

4. The soundfield transmitter would incorporate a simple to use mute mode for non-transmission of the teacher to teacher/learner conversations without the need to switch off the speaker system.
5. The system would be compatible with interactive whiteboards and other external audio devices.

The reason for excluding soundfield systems without the Datalogging facility is that previous study designs were susceptible to response and expectation bias as the teachers completed questionnaires on their use of the system but were not blind to the rationale of the study (Heeney, 2007, Dockrell and Shield, 2012). The reason for excluding soundfield systems with a fixed level of gain is that classrooms are environments with variable noise levels and a soundfield system that does not monitor the background noise and adapt the gain to the teachers' voice accordingly may not provide an appropriate SNR. Soundfield systems fitted with multiple room speakers were excluded due to the costs involved and health and safety issues. The school buildings selected for inclusion in this research were of an age where there was the potential for asbestos, and so the installation of this type would be prohibited. Having a simple to use mute function was included as previous studies indicated that complex soundfield systems resulted in a high attrition rate. Teachers in the classroom play audio from other interactive classroom technology and so any soundfield needed to allow this to occur seamlessly for consistent usage to be achieved.

This research used the Phonak Digimaster 5000 dynamic classroom soundfield system, which transmits on a digital 2.4 GHz band (fitting guidelines and manufacturers specifications are provided in Appendix 4). The portable system consists of a single speaker unit that houses 12 miniature speakers and an Insprio transmitter connected to a lapel microphone (Phonak, 2014). The system was selected for this research, as it is the only device on the market that collects and stores data on usage through the Datalogging system. Datalogging is commonly used in digital signal processing hearing aids and records information on the length of time a hearing aid has been worn since the last fitting. The Inspiro transmitter stores information on the use of the dynamic soundfield system every fifteen minutes. The system was installed to the manufacturer's specification; the speaker was installed at the front of the room, near the corner and mounted on the wall at a distance of 1.4m from the floor (Figure 11). A similar location is recommended for systems using a floor stand. Twelve classrooms used wall-mounted systems with one floor stand in a classroom without a suitable location for a wall bracket.



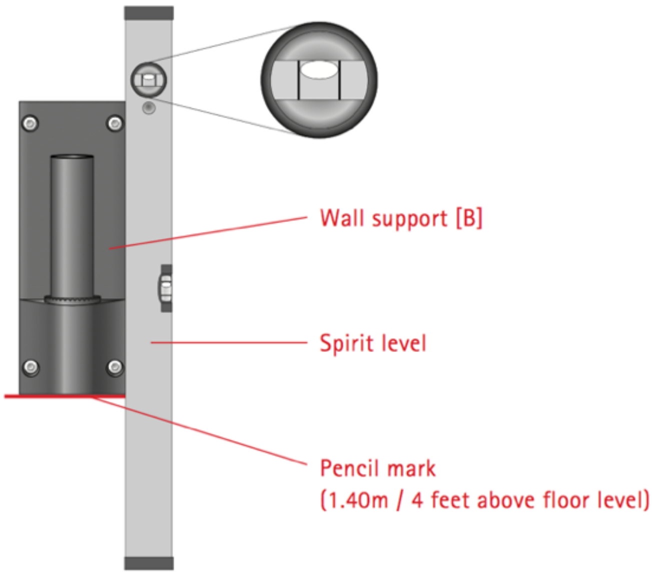


Figure 11: **Location and fitting guidelines of the dynamic soundfield system.**

3.6.2 Soundfield Amplification Training

Previous studies have found teaching staff either felt insufficiently trained in using the system (Dockrell and Shield, 2012, Heeney, 2007) or had been provided with training that promoted the benefits of soundfield and so introduced the potential for response and expectation bias (Stephenson, 2007). The training was provided to all teaching and support staff on an individual or whole school basis prior to the delivery of the equipment. The staff were given training and time to practise with the dynamic soundfield system so that they felt able to turn the system on and off, charge the system, position the microphone correctly and attach the audio cable from the soundfield system to the computer, which was connected to the interactive whiteboard. When the equipment was issued to the school supplementary training and a follow-up observation visit in class was provided within the first week. During the training and follow-up sessions, the teachers were given the opportunity to ask questions and the phone number and email address of the researcher was left. Additional questions were also raised at the monthly Datalogging collection sessions in schools. The instruction manual for the system was also provided to each school.

3.6.3 Soundfield Amplification Faults and Maintenance

In addition to insufficient training, faults and problems with the soundfield system all contributed to a high attrition rate in previous studies, which reduced the power of the sample and compromised internal validity (Anderson and Goldstein, 2004, Dockrell and Shield, 2012, Heeney, 2007, Massie and Dillon, 2006a, Massie and Dillon, 2006b). To ensure faults or problems with the equipment were addressed quickly, a supply of spares was retained centrally. These included lapel and transmitter belt clips, lanyards and lapel microphones. Also, two surplus dynamic soundfield systems were retained and these could be used to replace equipment that required more specialist repair. The phone number and email address of the researcher were left with the class and headteacher so that any faults could be reported directly. The researcher or staff from the Sensory Support (Deaf) Service provided onsite repairs and spares and the issuing of replacement equipment.

3.7 Pre-Intervention Measurements

3.7.1 Hearing Screening

Listening in the classroom can be compromised by changes in hearing thresholds and so prior to the intervention, all participants in the control and intervention classes had their hearing screened. Due to the number of participants and schools involved, the hearing screening was performed in the individual schools by a qualified educational audiologist or audiologist. Hearing loss exists where there is average raised air conduction thresholds of 20dB or greater in each ear over five frequencies: 250, 500, 1000, 2000 and 4000 Hz (BSA Professional Practice Committee, 2015). Air conduction thresholds were measured starting at 1000 Hz before testing 2000 Hz, 4000 Hz, 8000 Hz, 500 Hz, and 250 Hz as per British Society of Audiology pure tone audiometry procedures. Due to levels of ambient noise and timescale measuring hearing thresholds of 0dB was not an objective and a participant was deemed to have normal hearing thresholds when they tested below 20dB across the five test frequencies. Bone conduction was not tested on participants whose air conduction thresholds were within normal limits as this indicates there are no sensorineural issues. Furthermore, levels of ambient noise in the school made bone conduction testing unfeasible.

3.7.2 Equipment and Test Environment

A portable calibrated Otometrics Madsen Xeta audiometer was used in all the hearing screening assessments, which was checked to be functioning before each session (see Appendix 5 for the specification sheet). Excessive ambient noise levels can affect the test results as it masks the test signal especially at 500Hz and below (Lo and McPherson 2013). To attenuate the level of background and underlying noise TDH-39 supra-aural earphones with noise-excluding audiocups were used in each assessment. Furthermore, the choice of the test environment was controlled to ensure that background noise levels were not excessive. The test rooms were not near main roads, gym/lunch halls, toilets or had heating/cooling systems that could not be switched off. The type of rooms used were empty classrooms, medical rooms, an empty library, an unoccupied learning support room, and a meeting room. A Casella 620A Integrated Digital Sound Level Meter (Class 2) which conforms to standards ANSI S1.4, ANSI S1.43, IEC 61672, IEC 60651 and IEC 60804 was used to monitor the noise levels in the test environment (Refer to Appendix 6 for the specification sheet).

3.7.3 Hearing Screening Assessment

Each participant was tested individually, with the class teacher in the schools determining the order that each participant was assessed. Schools were booked for a hearing assessment in either a morning or afternoon and for learners that were absent follow up visits were arranged to ensure that all participants had a test for hearing acuity. Several variables including the motivation and mood of the participant affect the process of establishing hearing thresholds as does the quality of the instructions provided (Bamford et al., 2005). Prior to starting the test, it was explained that a sound like a beep or a whistle would be played and that every time it was heard even very soft sounds, the participant was to raise their hand. A practice session followed with the sound presented with the headphones on the table at a level of 90dB at 1000Hz and then with the headphones worn at a level of 40dB at 1000Hz. The student sat facing away from the audiometer at an angle that allowed the tester's hands on the audiometer not to be seen but the facial expressions of the participant to be viewed. As the threshold of an individual is a measurement based upon the relationship between physical stimulus using a behavioral response, seeing the learner's face helps validate the reliability of the response (Hesketh, 1986).

The '10dB down and 5dB up' rule was used to establish hearing thresholds. The participant was presented with a warble tone of between one and three seconds and the intensity of the stimulus was reduced in increments of 10dB until there was no response at which point the signal was increased in steps of 5dB until a response was recorded and then repeated at the same level twice (BSA Professional Practice Committee, 2015). Care was taken during the assessment that the timings of the stimulus were not predictable and that a variety of intervals between the tones was presented.

3.8 Room acoustical measurements

As discussed in Chapter 2, background noise in a classroom can mask speech which affects speech intelligibility. Long T_{mf} can also have a detrimental effect on speech intelligibility in the classroom, in contrast, rooms with shorter RT_{60} may enhance speech transmission. As these could be confounding variables, they were factored into the study design.

Acoustic surveys of all the control and intervention classrooms were undertaken. These included measures of ambient noise levels in unoccupied classrooms, room volume, long reverberation and early reflections. The data collection methods and results are discussed in Chapter 4. Noise surveys during lessons in numeracy, literacy and interdisciplinary learning (IDL) were also performed. Also, questionnaires were developed to measure the teacher and learners' views on the experience of noise levels in the classroom. The data collection methods and results are presented in Chapter 5.

3.9 Research instruments

3.9.1 Learners’ listening in noise questionnaires

In addition to noise surveys, the views of the learners were sought to establish if there was a different experience of noise between those exposed to dynamic soundfield amplification and the control group. The method used to capture the learners’ experience of noise in the classrooms was a questionnaire. The reasons for using a questionnaire were twofold. Firstly, research suggests that noise causes annoyance, fatigue, and stress to the occupants of classrooms and these emotions cannot be measured directly (Wålinder et al., 2007, Tiesler and Oberdörster, 2008). Secondly, although the perceived loudness of a sound is the primary psychological correlate of physical intensity, judgements on loudness depend on other factors such as the distance the individual is from the source, context in which the sound is heard, and the nature of the sound (Moore, 2003, Hellbrück et al., 2014).

For a questionnaire to have validity it measures the issue it is attempting to measure and should be constructed in a way that the participants understand the objective of the questions (Field, 2013). To this end, the aim of the printed noise questionnaire was fourfold: Firstly, to establish if there was a significant difference between learners’ awareness of different noises commonly heard in classrooms in both the control and intervention classrooms. Secondly, to determine if the learners’ perceptions of noise levels in nine different curricular areas were statistically different between the two conditions. The third objective was to gauge the feelings of the learners in the different conditions towards noise using seven complementary antonyms. Finally, to establish if learners exposed to dynamic soundfield could hear the class teacher more easily compared to the control. Six different classroom scenarios were presented.

Rea (2014) observes that the researcher requires to be cognisant of the participants being sampled in terms of the language used in the questions and length of the questionnaire. As the age range of the learners was young, the questions and response method required to be accessible as some participants may have difficulty with literacy. A symbolised closed questionnaire was developed for the learners using Picture Communication Symbols. The questions used symbols that were selected on iconicity, which is defined as a perceived relationship between the symbol and its referent and so establishes a visual similarity (Dada et al., 2013). Fife Council has been symbolising the whole primary school environment for several years to support learners' communication and access to the curriculum. Fife Assessment Centre for Communication through Technology (FACCT) has been central to this project and so links were made to ensure that the symbols used were in keeping with the format familiar to the learners.

The participants were asked to respond to the questions by drawing a circle around the answer which was either in a form of a list or makes a judgement about a situation using a Likert scale (see Appendix 7). This format was selected as it minimised the risk of incomplete or missing data which would compromise the statistical power and produce biased estimates. The listening in noise questionnaire was not piloted before use. The data collection methods and results are presented in Chapter 5.

3.9.2 Teacher questionnaire

All teachers in the control and intervention classrooms were asked to complete a questionnaire (See Appendix 8). The survey was designed to collect the following data:

1. The number of years of teaching experience.

2. The teacher's awareness of both external and internal noise sources in the classroom.
3. The level of noise during nine different curricular areas.

One reason for collecting data on teaching experience was to establish if this was equally distributed between the two groups. The purpose of asking about the sources and levels of noise was twofold: to explore if the teachers' experience of noise was different between the control and intervention classrooms. Also, to compare whether there were similarities between the teachers and young learners' experiences of noise.

In addition, the teachers in the intervention classroom were asked about:

1. The impact of the dynamic soundfield system on different aspects of learning using a five-point Likert scale.
2. Their feelings about wearing the dynamic soundfield system.
3. Whether background and underlying noise levels were reduced in classrooms exposed to dynamic soundfield.

At the end of a longitudinal study, it seemed appropriate to seek the views of teaching staff using the equipment. The results discussed in Chapter 5 should, however, be treated with caution as the teachers were not blind to the intervention and so the responses may be susceptible to bias.

3.9.3 Performance Indicators in Primary School (PIPS)

All young people entering school in Fife complete the Performance Indicators in Primary School (PIPS) baseline and follow-up assessments. The test is generally administered by a member of the teaching staff on an individual basis. The assessments are composed of two parts: the first section measures cognitive development (early mathematics, early reading, picture vocabulary, and phonological awareness); the second part measures social and personal development. Data was not collected on social and personal development as this was not relevant to the study design. Fife Council also did not provide separate data on picture vocabulary.

The Early Mathematics module consists of a range of assessments. The counting assessment involves four objects appearing on the computer, the learner counts the objects which then disappear, the learner is then asked to recall the number of objects they saw. This is repeated with seven objects. Digit identification assessments include single, double and three-digit numbers. The assessment also includes informal simple addition and subtraction sums without symbols. There are also formal mathematical problems presented with notation. Finally, the vocabulary associated with mathematical concepts is assessed (Scottish Government, 2016a).

The reading assessment involves a vocabulary test in which the young learner is asked to identify a number of objects embedded in a picture on the computer screen. The learner is also asked to identify letters and write their name. The learner is asked to recognise words and read aloud simple sentences. The words are common and high-frequency words. The reading task has two more advanced assessments: Walking to School and Cats. The young learner reads the sentence and is then asked at certain points to select one word from three (one target, two foil) that is most appropriate to the context. In the phonological awareness assessment, the learner is asked to repeat words and nonsense words they hear. The learner is also presented with a three-picture matrix (one target, two foil) and is asked to select a word that rhymes with a target word (Scottish Government, 2016a).

3.9.4 Achievement for Excellence standardised tests

Improvements in educational attainment in primary 3 were measured using the AfE (InCAS) suite of assessments. AfE (InCAS) is a standardised, norm-referenced, computer-based school-administered assessment. It was selected for this research as it was not only used widely in Scottish schools but also internationally and so provides a standardised measure of academic achievement and progress (Tymms et al., 2000, Tymms et al., 2004).

As Table 2 illustrates the AfE (InCAS) assessment is composed of six modules: Reading, General Mathematics, Mental Arithmetic, Developed Ability, Spelling, and Attitudes. Attitudes gather information using a sliding scale (positive through to negative) on the learners' feelings towards reading, mathematics, and school. As the attitude of the learners was not relevant to the current study design, data on this module was collected, as is part of the suite of assessments but it was not analysed.

Reading and Spelling	*Maths*	*Attitude and Developed Ability*
Word Recognition *- The learner hears a word, which includes a contextual sentence. The learner then selects the target word from a choice of five on screen.*	**General Maths 1** - This cover counting, informal arithmetic, partitioning and place value, fractions and decimals.	**Attitudes -** Attitudes towards Reading, Maths and School are assessed using a sliding scale.
Word Decoding *- The learner hears a nonsense word and they select the target word from a choice of five words on screen.*	**General Maths 2 -** This covers sorting, patterns, formal arithmetic, problem solving and algebra. The questions are non-curriculum based. Measures, Shape and Space. Data Handling.	**Picture Vocabulary -** The learner hears a word and selects the picture that best represents that word.
Comprehension *- The learner reads through a passage and, when given a choice of three words, must select the word that fits into the sentence. most appropriately.*	**Mental Arithmetic -** Assesses the learner's ability to process numerical operations quickly and accurately.	**Non-Verbal Ability -** A pattern appears on the left-hand side of the screen and the learner must then find the corresponding pattern within a larger pattern on the right-hand side.
Spelling – *the learner hears a word, in a contextual sentence. They then use the onscreen keyboard to select the correct letters for the target word.*		

Table 2: **Different modules and subtests from the AfE (InCAS) computer-based, adaptive assessments (Centre for Evaluating and Monitoring, 2014).**

Each participant was provided by the CEM with a unique computer login code and password that was discrete to each assessment. The participants worked individually on a computer, with the automated instructions and questions being presented via individual headphones. The participants were tested in small groups with the size of the group determined by the number of available computers. All the testing was completed in a separate room to minimise distractors. The pre-intervention assessments were either invigilated by teaching staff in the school (N=2) or by teachers and support staff from the Sensory Support (Deaf Learners) Service (N=11). The post-intervention assessments were invigilated by the school staff as the tests are routinely completed by these institutions at the end of the academic year.

The teaching staff not only invigilated the assessment but also allocated the unique codes for each unit to the learners. The teachers also checked all the equipment was functioning. Those invigilating the assessments were not blind to which was the control or intervention classrooms. However, they did not deliver or grade the assessments as this was completed by CEM at the University of Durham who were blind to which participants were in the control and intervention classrooms. As Table 3 illustrates, Reading, Mathematics and Developed Ability units last a maximum of twenty-five minutes. Spelling and Mental Arithmetic have a maximum time limit of twenty minutes. The assessments were timed and could not be paused. The guidance recommends that no more than two modules should be completed each day and the process should begin with the Attitudes module. All other assessments can be presented in any order. Apart from the Attitudes module, all other units were tested in a random order. The assessments were completed in the schools over a three-week period. Repeat visits were made to absent participants. This ensured all participants completed the assessments. The data analysis and results are discussed in Chapter 6.

Timing of AfE (InCAS) assessments		
Reading – 20 to 25 minutes	**General Maths** – 20 to 25 minutes	**Developed Ability** – 20 to 25 minutes
Spelling – 15 to 20 minutes	**Mental Arithmetic** – 15 to 20 minutes	**Attitudes** – 5 to 10 minutes

Table 3: Anticipated timings for each of the AfE modules.

3.10 Measuring achievement

3.10.1 Developed Ability

The Developed Ability module is composed of two subtests: Picture Vocabulary and Non-Verbal Ability. The picture vocabulary subtest involves the learner hearing a word, being presented with a five-picture matrix on the computer screen (one target and four foil) and then selecting the picture that best represents that word within the matrix. The adaptive assessment measures the learner's understanding of vocabulary that increases in complexity, depending on ability. In the Non-Verbal Ability subtest, a collection of shapes is presented on one side of the computer screen and the learner is required to find a corresponding pattern within a larger pattern on the other side of the screen. The adaptive assessment is measuring the learner's ability to use reason to solve difficult problems (Centre for Evaluating and Monitoring, 2014). For both of the cognitive assessments, an adaptive algorithm matches the difficulty of the assessment to the ability of the learner (Tymms and Hanna, 2008). The Screenshots from all the AfE (InCAS) assessments are presented in Appendix 9.

The combined Picture Vocabulary and Non-Verbal Ability subtests are used to calculate an overall Developed Ability score (DAS) using the following formula (Centre for Evaluating and Monitoring, 2012):

$$\text{DAS}= \left(0.5905\ \times\ PV_{age}\right) + \left(0.4312\ \times\ NV_{age}\right) - 0.1259$$

$$\text{Error} = \frac{1.579}{\sqrt{\left(\dfrac{1.6468}{PV\_Age\_error}\right)^2 + \left(\dfrac{1.3414}{NV\_Age\_error}\right)^2}}$$

3.10.2 Reading and Spelling

The reading and spelling suite of assessments consists of four subtests: Word Recognition, Word Decoding, Reading Comprehension and Spelling. The Word Recognition subtest requires the learner to listen to a word, a sentence is then presented putting the word into context and the learner then identifies the written word from a five-word matrix (one target and four foil words). The adaptive assessment measures the ability of the learner to identify increasingly difficult written words. The Word Decoding subtest involves a five-word matrix of nonsense words presented on the computer screen (one target and four foil words), the learner hears a nonsense word and is required to identify the correct written representation of the word. The adaptive assessment measures the ability of the learner to decode increasingly difficult print. The Reading Comprehension subtest requires the learner to read a passage. Approximately, every fifth word involves the learner selecting the correct word from a list of three (one target and two foil) that is most appropriate to the context of the sentence. Both cognitive and metacognitive skills are required to complete this task (Merrell and Tymms, 2004). The adaptive assessment is measuring the ability of the learner to understand the meaning behind the printed word. The spelling subtest requires the learner to spell the word presented using the online keyboard. The adaptive assessment measures the ability of the learner to spell words that are increasingly more difficult (Centre for Evaluating and Monitoring, 2014).

The combined Word Recognition, Word Decoding, and Reading Comprehension subtests are used to calculate an overall Reading score (ORS) using the following formula (Centre for Evaluating and Monitoring, 2012):

$$ORS = (0.3169 \times WR_{age}) + (0.1704 \times WD_{age}) + (0.4992 \times Comp_{age}) + 0.145$$

The error rate is calculated using the following formula when all the learner completes all three subtests:

$$\text{Error} = \frac{1.2273}{\sqrt{\left(\frac{1.0365}{WR\_Age\_error}\right)^2 + \left(\frac{1.4133}{WD\_Age\_error}\right)^2 + \left(\frac{1.2256}{Comp\_age\_error}\right)^2}}$$

Access to the Reading Comprehension subtest is conditional on the scores achieved in the other two subtests. If the combined Word Recognition and Word Decoding age equivalent scores are less than 8.0 years, the learner will not be presented in the Reading Comprehension subtest. Learners not meeting the threshold to access the reading comprehension (RC) subtest have their reading age equivalent score calculated using the following formula:

$$RC = \left(0.3169\ \times WR_{age}\right) + \left(0.1704\ \times\ WD_{age}\right) + (0.4992\ \times 4.0) + 0.145$$

The error rate is calculated as follows:

$$\text{Error} = \frac{1.2273}{\sqrt{\left(\frac{1.0365}{WR\_Age\_error}\right)^2 + \left(\frac{1.4133}{WD\_Age\_error}\right)^2}}$$

A Reading Comprehension age equivalent score of less than four is therefore replaced by four as a learner achieving a very low score on the comprehension test could theoretically score less than a learner not presented in the reading comprehension assessment. Therefore, the lowest age equivalent score in the comprehension module is 4. Any learners not making the threshold for the assessment had their age equivalent score replaced by 4 (Centre for Evaluating and Monitoring, 2012).

3.10.3 General Mathematics

The General Mathematics suite of assessments consists of four subtests: Number 1, Number 2, Data Handling and Measure, Space and Shape (MSS). All four tests involve the learner listening to a mathematical question and selecting the answer on the computer screen. The Number 1 subtest assesses counting, informal arithmetic, partitioning, place value, decimals, and fractions. The Number 2 subtest assesses formal arithmetic, patterns, sorting, problem-solving and algebra. The Data Handling subtest assesses the ability of the learner to interpret a range of data in different formats. The MSS subtest assesses the ability of the learner to measure (length, weight, volume/capacity, time and temperature) and explore a variety of shapes. There is no published or unpublished information on the methods used to calculate the overall General Mathematics score.

3.10.4 Mental Arithmetic

The Mental Arithmetic suite of assessments consists of four subtests: addition, subtraction, multiplication, and division. All learners access the addition subtest. Thereafter, learners are only presented for the subtraction subtest if the addition age equivalent score is >5 years. Learners only access the multiplication subtest if the subtraction age equivalent score is >7 years. Finally, the division subtest is only sat if the multiplication age equivalent score is >8 years. The number of participants presented for each subtest decreases as the threshold for accessing the tests increases. Following advice from the AfE (InCAS) coordinator in Fife Council, it was decided not to include the individual subtests in any analysis due to the high attrition rate and instead measure the overall Mental Arithmetic module scores. There is no published or unpublished information on the methods used to calculate the overall General Mathematics score.

3.10.5 Age equivalent and standardised scores

The computerised algorithm within AfE (InCAS) produces a feedback profile that includes the age of the learner, age equivalent scores, age difference scores, associated error values, and standardised scores. The age of the learner is calculated for each assessment using the individual’s date of birth in conjunction with the date that each unit was assessed. A general methodology is used to calculate age equivalent scores by CEM, although there are minor variations between each module. Rasch scaling is used to create equal interval scales. For many of the modules, the algorithm takes the age of the learner as the starting point and selects items that a young person three years younger would have fifty per cent chance of answering correctly. The questions get progressively more challenging until a predetermined set of incorrect answers results in the program estimating the age of the learner for that module. A second and third group of presentations are then made to refine the original estimated age equivalent score (Merrell and Tymms, 2004). Age difference scores are calculated by arithmetically subtracting the age of the learner on the date when the test was completed from the age equivalent scores achieved in the test.

The Reading Comprehension module is slightly different as the passage presented to the learner is not determined by the participant's age but by the scores achieved in the Word Recognition and Word Decoding subtests. Learners are only presented into the Reading Comprehension test if their age equivalent scores in the Word Decoding and Word Recognition subtests add up to 8 years or more.

Age standardised scores are provided for the four modules: Developed Ability, General Mathematics, General Arithmetic, and Reading. Age equivalent values are provided for the associated subtests. Age standardised scores (AS) are calculated using the following formula (Centre for Evaluating and Monitoring, 2012):

$$AS = \left(\frac{score\ - Age\ at\ Test\ - Mean}{Standard\ Deviation}\right) \times 15 + 100$$

Improvements in educational outcomes were measured using both age equivalent difference values and standardised scores. Age equivalent difference values were collected for each subtest and standardised scores for each of the four modules.

3.10.6 Summary

This chapter has presented the study design developed to answer the research questions on whether dynamic soundfield amplification delivers improvement in educational attainment and reduces the poverty-associated attainment gap between learners from SIMD 1 and 5. A description of the independent variable, dynamic soundfield system and its installation, training and Datalogging collection methods have been presented. Furthermore, the dependent variable, standardised measure of attainment and the methods used to collect the data has been addressed. The selection of participants, school location, noise surveys and reverberation measures that form a central core of the study design have also been addressed. The method to control confounding variables, such as screening hearing has also formed part of this chapter.

The next chapter examines the characteristics of the participants with the view of establishing internal and external validity. Data collected from the acoustic measurements of the classrooms and usage of the soundfield system will be discussed.

Chapter 4
Internal and External Validity

4.1 Aims of the Chapter

This chapter will address secondary research aims two and three of this thesis (see section 3.2.2) by establishing internal and external validity. External validity in quantitative research is regarded as a major criterion for evaluating the quality of a study (Slack and Draugalis, 2001). External validity is an issue about generalisations, the extent to which the outcomes of the study can be applied to and across other populations or settings (Campbell and Stanley, 1966, Johnson and Christensen, 2000, Slack and Draugalis, 2001). The more representative the sample participants are the more confident that the outcome can be extrapolated to other identified groups. Polit and Beck (2010) argue that for external validity to be established there needs to be sufficient descriptive detail about the characteristics of the participants and the setting from which the data was collected. Furthermore, Lucas (2003) observes that in many experimental studies it does not specify a target population for generalisation. Sections 4.2, 4.3 and 4.4 compare the characteristics of the participants, teachers, and schools in this research to two other groups: all primary school learners and teachers across Scotland and all learners at the primary 3 stage. The purpose is to demonstrate sufficient similarities and so allow generalisation of results.

Campbell and Stanley (1966) observe that for a study to establish the cause and effect relationship between an intervention and the observed outcome it needs to demonstrate that extraneous factors are unlikely to have influenced the results. As discussed in Chapter 3, ambient noise levels in unoccupied classrooms along with the early to long sound energy ratios were identified as situational variables that could compromise internal validity. Section 4.5 of this chapter will present the data on the acoustic properties of the control and intervention classrooms.

Use of the soundfield systems in the intervention classroom was also identified as an extraneous variable as previous studies have been compromised through inconsistent or non-usage of the soundfield systems which increases the attrition rate and reduces statistical power (Massie and Dillon, 2006a, Dockrell and Shield, 2012). Furthermore, the method used to established use of the systems has been through questionnaires that were subject to response and expectation bias as the teachers were not blind to the intervention (Heeney, 2007, Massie and Dillon, 2006a). In the present study data on the daily use of the soundfield systems was collected through the inbuilt Datalogging firmware which stores information in 15-minute timeframes. Section 4.6 will present and analyse the quantitative data collected from each of the dynamic soundfields.

Speech intelligibility not only depends on the acoustics of the classroom but also the level of background noise relative to speech and so was also identified as an additional variable. Noise surveys were also completed in each of the control and intervention classrooms during lessons on literacy, numeracy, and IDL. In addition, the learner's and teacher's views on the perceived levels of noise were collected through questionnaires. Due to the amount of data collected this will be discussed in Chapter 5.

4.2 Participant Attrition

The participants were recruited prior to the allocation of control and intervention status. This method was used to minimise the effect of selection bias which could compromise internal validity. There was concealed randomisation as the headteacher in each school allocated the learners to the individual classes blind to which would be the control and intervention classrooms. Torgerson (2008) advises that if an intervention is to be delivered by a class teacher then the teacher needs to be linked to the class prior to the allocation. All teachers were linked to a classroom before the allocation of intervention and control status.

An opt-out consent letter was sent to each participants' family (n=515) providing details about the research. Torgerson et al. (2013) observe that in experimental designs involving control and intervention groups bias and attrition may occur if a class or participant is randomly allocated to a control group when they anticipated gaining the benefit from an intervention. This may lead the researcher to re-allocate a class to an intervention or for participants to drop out of the study leading to attrition. As part of the University of Edinburgh ethics application, it was agreed that any class not fitted with a dynamic soundfield during the data collection year would be provided with one the following year. By providing intervention at a future date it was hoped to address not only ethical issues but also factors of attrition and bias.

Only three families decided to withdraw from the research at the consent stage. At the pre-intervention stage, seven participants were identified as having raised hearing thresholds and, in each case, the family was contacted by phone on the day and a referral made to the local audiology department. Five were bilateral referrals and two single-sided (left ear) referrals. Although the hearing screening assessments ensure that the participants in the control and intervention groups have hearing within normal limits, young people can be susceptible to various middle ear complaints such as Otitis Media with Effusion (OME) which means that at different points in the year their hearing may fluctuate. Around 80% of young people will have had a least one episode of OME by the age of 10 years. The incidence increases in the winter months. The mean duration of effusions is 6 to 10 weeks but some cases are more persistent (NICE, 2014).

Having normal hearing thresholds was an inclusion criterion for participants and so the data on the seven participants with reduced hearing thresholds were excluded from the sample size. Once the parents had consented to their child being involved in the research each learner also completed a consent form. In total, 505 young people and their parents were recruited to participate in this research study and consent was provided.

As Figure 12 illustrates there was a marginal amount of attrition during the study. Between the pre-intervention and post-intervention assessments, ten learners left the research schools. This accounted for 1.98% of the sample size with a marginally higher dropout rate in the intervention classrooms (2.11%) compared to the control (1.81%). As Amico (2009) rightly observes attrition in some form is a reality when undertaking longitudinal studies with people. Dealing with attrition data is primarily discussed in the literature on random control trials and can affect external validity by producing a final sample that is not representative of the target population or introduces bias and a reduction of statistical power (Amico, 2009, Torgerson et al., 2013). Schulz and Grimes (2002) observe that loss of follow-up data of less than five per cent is of little concern as the amount of bias will be minimal. As the attrition is small the data of the ten participants were removed from the final sample size.

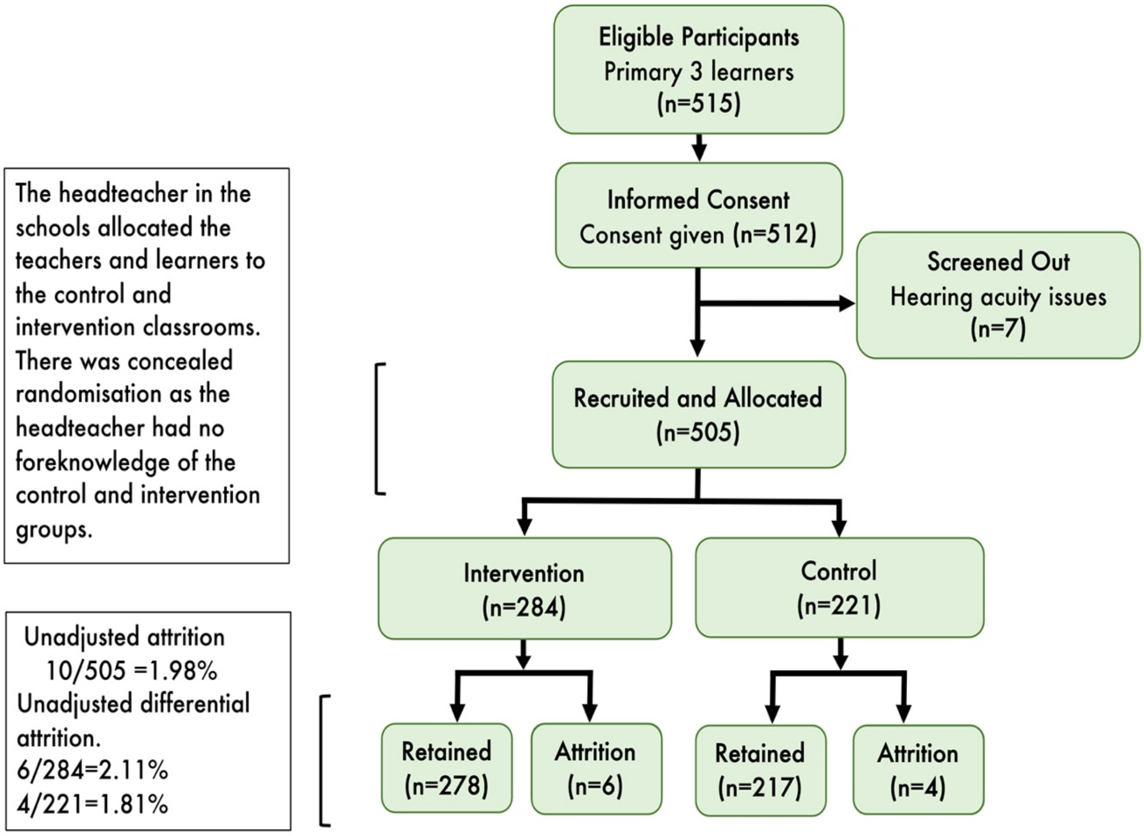


Figure 12: Participants' allocation and associated attrition in both the control and intervention groups.

There is a consensus that reporting on the attrition data is often overlooked and the Consolidated Standards of Reporting Trials (CONSORT) advocates publishing the data in studies (Consolidated Standards of Reporting Trials, 2010, Amico, 2009, Schulz and Grimes, 2002). In keeping with the CONSORT guidelines Table 4 provides data on the attrition rates, reasons for exclusion along with some characteristics of ten participants that moved schools during this research.

	All Learners	*Control*	*Intervention*
pre-intervention assessment	505	221	284
post-intervention assessment	495	217	278
did not receive post-intervention assessment	10	4	6
reason	10 moved schools		
analysed	495	217	278
excluded from analysis	10	4	6
reason	Small proportion of the sample (1.98%)		
gender	Male=6 Female=4	Male=2 Female=2	Male=4 Female=2
SIMD	SIMD 1=4 SIMD 2=3 SIMD 4=1 SIMD 5=2	SIMD 1=1 SIMD 2=1 SIMD 5=2	SIMD 1=3 SIMD 2=2 SIMD 4=1
age (mean) at pre-intervention stage	7:1	7:1	7:3

Table 4: Study attrition rates and data excluded from analysis.

Excluding the attrition data, 495 participants were involved in the study. Based on the 2015 pupil census this sample represents 0.86% of the primary 3 population (N=57,750) of Scotland.

4.3 Characteristics of Participants

In Scotland, every learner is allocated a Scottish Candidate Number, and this was used to extract the participants' age, gender, ethnicity, language and additional support needs from SEEMiS. Data on the target population was primarily based on the associated files from the publicly funded annual pupil census (Scottish Government, 2015a). Each September the thirty-two Scottish local authorities submit their SEEMiS data to ScotXed, the education analytical division of the Scottish Government, which becomes the pupil census and is published in December of the same year (Scottish Government, 2015a). The pupil census data in this research is based on the 2015 submission.

Although any central data management system is prone to human error, the SEEMiS system was used to collect participants' data for two reasons. Firstly, at the start of the academic term schools provide parents with a printed copy of the SEEMiS record held on their child and are asked to check and amend the file. If parents were asked to complete an additional form seeking similar information then this could be perceived as duplication, resulting in a high non-return rate. Furthermore, in establishing external validity it is appropriate that the data on the research participants and target population are taken from the same SEEMiS database source.

4.3.1 Gender Ratio of learners

As Table 5 illustrates, 51.9 per cent of the participants were male and 48.1 per cent were female (n=495). This is comparable to the national figures for the whole primary school sector (n=391,148) which has a male to female ratio of 51% and 49% respectively. The figure for those at the primary 3 stages across Scotland (n=57,750) is also 51% and 49%. Although this representation is maintained in the control and intervention groups, the ratio between males and females is marginally higher in the control group. A Chi-square test indicated that there was no statistically significant difference in the gender composition between the control and intervention groups, χ^2 (1, N=495) =0.059, p=0.809.

	Research Learners	***Control***	***Intervention***
n	495	217	278
Male (n, %)	257 (51.9%)	114 (52.5%)	143 (51.4%)
Female (n, %)	238 (48.1%)	103 (47.5%)	135 (48.6%)

Table 5: Gender ratio of the participants.

Figure 13 provides the gender ratio of participants based on the SIMD rating. Two of the main assumptions of the Chi-square test of independence is that each participant can only contribute data to one cell in any contingency table and that each cell should have a minimum value of five (Heiman, 2013). The SIMD 3 quintile did not meet this assumption and so was not analysed. The results found there was no significant difference between the control and intervention groups on gender by SIMD quintile: SIMD 1 χ^2 (1, N=155) =0.19, p=0.892, SIMD 2 χ^2 (1, N=76) =0.492, p=0.483, SIMD 4 χ^2 (1, N=55) =3.51, p=0.06 and SIMD 5 χ^2 (1, N=183) =1.63, p=0.202.

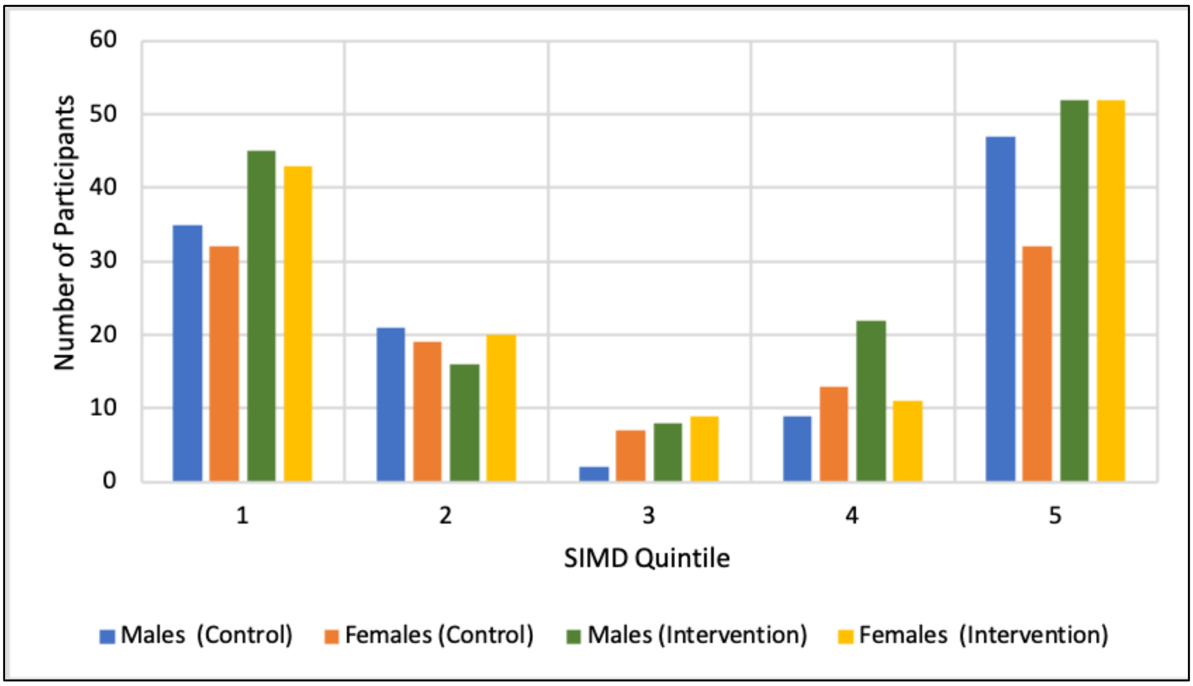


Figure 13: Gender by SIMD quintile in the control and intervention classrooms.

4.3.2 Learner ethnicity

The number of ethnic minority learners in Scotland is small and this is reflected in the sample group composition. To facilitate a comparison the two largest ethnicity groups, White/Scottish Other and Asian were combined with other similar categories. White/Scottish Other included White Scottish, Other, Gypsy Traveller, Irish and white Polish. Asians included Indian, Pakistani, Bangladeshi and Chinese. This is similar to the methods used by Gillooly and Riddell (2019) when presenting an overview of the statistics on ASN in Scotland and special educational needs in England. As Table 6 illustrates using the 2015 pupil census the proportion of learners in each ethnic grouping reflects the ethnic distribution of students attending public-funded primary schools across Scotland. Approximately, 93% (n=461) of the learners in this research were classified as White Scottish/ Other, which is slightly higher than the national average, but this may reflect that the Pupil Census had a 2.33% not disclosed submission rate (n=9118). This ethnic distribution was maintained in both the control and intervention groups. A Chi-square test confirmed that there is no significant difference between the control and intervention groups on ethnicity, χ^2 (1, N=495) =0.105, p=0.746.

Ethnicity (n, %)	***All Scottish Primary Learners n=391,148***	***All Scottish Primary 3 Learners n=57,750***	***All Research Learners n=495***	***Control n=217***	***Intervention n=278***
White Scottish/ other	35,3949 (90.5%)	52,239 (90.5%)	461 (93.1%)	204 (94%)	257 (92.4%)
Mixed	5117 (1.3%)	774 (1.4%)	3 (0.6%)	1 (0.5%)	2 (0.7%)
Asian	15536 (4%)	2345 (4%)	17 (3.4%)	5 (2.3%)	12 (4.3%)
Caribbean/ Black	605 (0.2%)	94 (0.2%)	0	0	0
African	3865 (1%)	624 (1%)	3 (0.6%)	0	3 (1.1%)
Arab	1169 (0.3%)	193 (0.3%)	2 (0.4%)	0	2 (0.7%)
Other	1789 (0.5%)	236 (0.4%)	4 (0.8%)	4 (1.8%)	0
Not disclosed	9118 (2.3%)	1245 (2.2%)	5 (1%)	3 (1.4%)	2 (0.7%)

Table 6: Percentage of learners within each ethnic group based on the 2015 pupil census and SEEMiS database of the same year (Scottish Government, 2015a).

4.3.3 Language of the learners

Previous literature (Crandell and Smaldino, 1996, Nelson et al., 2005) had identified the influence of variables such as having English as an additional language when listening in noise. Therefore, information on the first language and competency level of all the participants was harvested from the Pupil Census using the individual's Scottish Candidate Number. The pupil census does not provide data on first language and fluency levels in primary schools and therefore there is no target population to establish external validity. The only published data was for learners across all Scottish schools (primary and secondary) and the 2015 pupil census recorded this at 5.2% (Scottish Government, 2015a). This is comparable to the current sample where 3.83% (N=19) of the participants had English as an additional language. Urdu, Punjabi, Cantonese, Akan, and Arabic were chronicled with Polish being the most common.

A histogram of the language competency levels of the learners that have English as an additional language in the control and intervention classrooms is shown in Figure 14. The number of learners in both conditions is small. There is a slightly higher number of participants in the intervention classroom that are classified as fluent in English compared to the control. Overall, the majority of participants have spoken English as their first language. A two-tailed independent t-test (α=0.05) confirmed there was not a significant mean difference between the control and intervention groups on the different language competencies of the participants: Developing Competency, $t(493)=0.35$, $p=0.724$, Competent, $t(493)=0.79$, $p=0.425$ and Fluent, $t(493)=-1.08$, $p=0.28$.

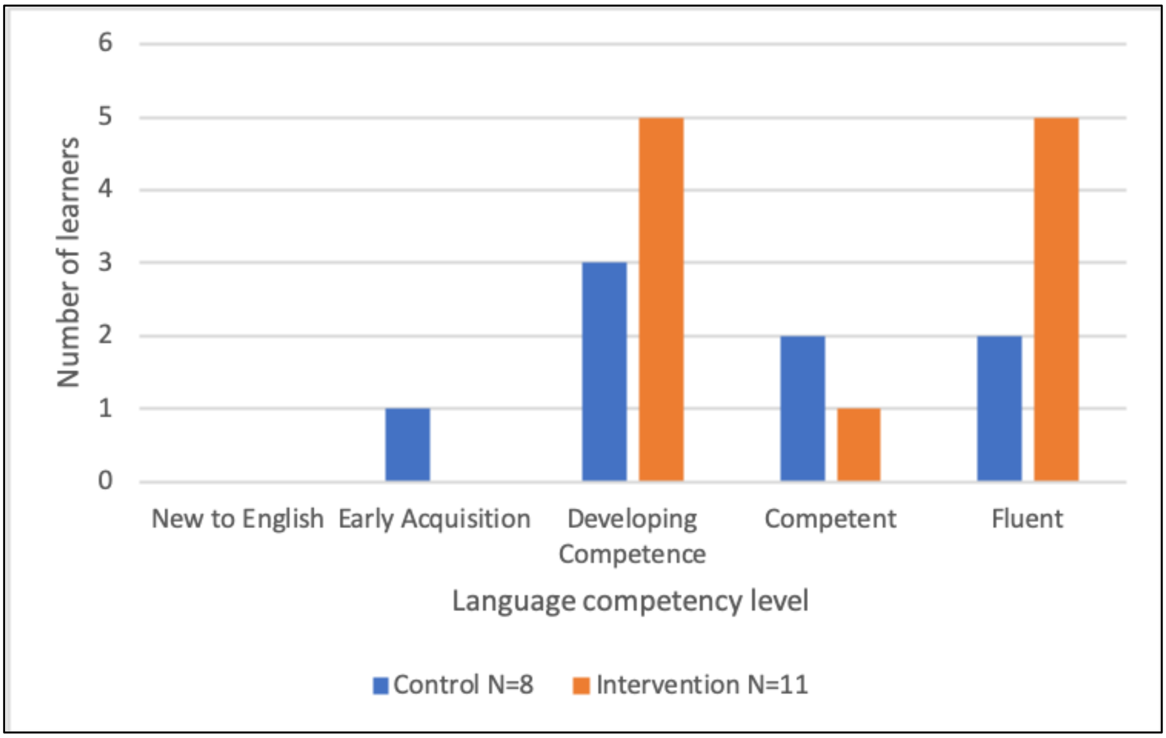


Figure 14: Language competency levels of learners in this study that have English as an additional language in the control and intervention classrooms extracted from the SEEMiS *database*.

4.3.4 Additional Support Needs of the learners

The Education (Additional Support for Learning) (Scotland) Act 2004 (as amended 2009) created the term additional support needs (ASN). Any form of barrier to learning is deemed an ASN and this can arise from factors such as social, emotional, cognitive, linguistic, disability or family and care circumstances (Scottish Government, 2009). As Table 7 shows, in 2015 there were twenty-three different ASN categories within the primary education sector. Eighteen different additional support needs were identified amongst the participants, five fewer than those provided by the national census. The red font indicates the ASN excluded from the study design. Hearing impairment and Deafblind were excluded as normal hearing acuity was one of the inclusion criteria. The green font indicates ASN not represented in the sample. Young carer, bereavement and substance misuse were not recorded as an ASN amongst any of the participants however, this only represents 0.37%, 0.8%, and 0.08% of the total ASN population in Scotland. The prevalence of ASN is higher for boys than girls in both the sample population and the national level.

Nature of Additional Support Need	Participants (%)		All Scottish Primary Learners (%)	
	Male	Female	Male	Female
Language or speech disorder	63.64	36.36	70.79	29.21
Social, emotional and behavioural difficulty	70.00	30.00	75.20	24.80
English as an additional language	55.56	44.44	52.27	47.09
More able pupil	66.67	33.33	55.03	44.97
Learning disability	62.50	37.50	64.71	35.29
Other moderate learning difficulty	50.00	50.00	61.01	38.99
Physical or motor impairment	85.71	14.29	69.14	30.86
Family Issues	42.86	57.14	53.93	46.07
Dyslexia	80.00	20.00	62.93	37.07
Other specific learning difficulty (e.g. numeric)	80.00	20.00	59.56	40.44
Visual impairment	0.00	100.00	58.34	41.66
Looked after	75.00	25.00	53.07	46.93
Autistic spectrum disorder	100.00	0.00	84.30	15.70
Physical health problem	25.00	75.00	56.36	43.64
Communication Support Needs	100.00	0.00	73.04	26.96
Mental health problem	100.00	0.00	71.53	28.47
Interrupted learning	0.00	100.00	55.62	44.38
Risk of Exclusion	0.00	100.00	92.02	7.98
Hearing impairment	0.00	0.00	56.79	43.21
Deafblind	0.00	0.00	60.87	39.13
Young Carer	0.00	0.00	46.51	53.49
Bereavement	0.00	0.00	56.00	44.00
Substance Misuse	0.00	0.00	59.55	40.45

Table 7: Percentage of learners in this study identified as having an ASN by gender compared to the whole Scottish primary population based on the 2015 pupil census and SEEMiS database record of the same year (Scottish Government, 2015a).

20.4 per cent of learners within the primary school system in Scotland (n=391,148) had an ASN in 2015. Previous literature has highlighted that the prevalence of ASN increases with age until a peak of approximately nine years old (Bradshaw et al., 2012). Therefore, a lower figure than the national average would be expected when measuring younger people. *Growing up in Scotland*, a longitudinal study tracking the lives of learners and their families collected data on 8-year olds based on 3657 parent interviews. At the time of the interviews, 60 per cent of the learners were in primary 3 and 40 per cent in primary four. It recorded that 13 per cent of learners had an ASN with 31 per cent having more than one type of need (ScotCen, 2015). This is comparable to the participants in this research where 14 per cent (n=70) had an additional support need with 27.14 per cent (n=19) having more than one type of need. Overall, 100 separate additional support needs were recorded amongst the 70 learners. Fifty-four point three per cent (n=38) were located in the intervention classrooms and 45.7% (n=32) were in the control. A comparison between the control and intervention classrooms on the number of learners with an ASN using a Chi-Square test indicated there was not a significant difference: χ^2 (1, N=495) =0.117, p=0.733.

The most commonly recorded ASN amongst 8-year-olds were language and communication issues, learning disabilities and social or behavioural problems (ScotCen, 2015). The most commonly recorded additional support needs amongst the learners in this research were language and speech disorder, social, emotional and behavioural difficulty and English as an additional language. The prevalence and range of ASN amongst the participants are in keeping with national data for younger learners. Furthermore, the range of needs is evenly spread across both the control and intervention groups.

Figure 15 provides the spread of ASN across the two conditions based on SIMD quintiles. There is a fairly even distribution across both control and intervention classrooms.

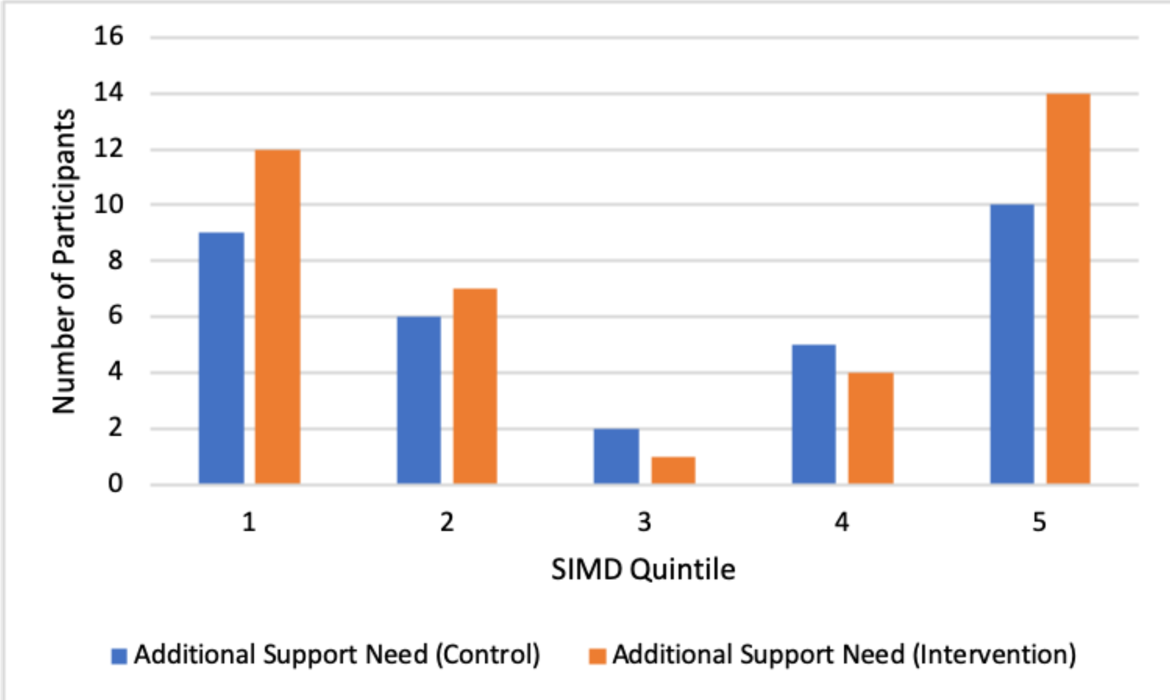


Figure 15: Distribution of additional support needs across the control and intervention classrooms based on the SIMD quintile.

4.3.5 Age range of the learners

All learners in the study spanned the age range of students found in the primary three classrooms in Scotland. There are usually seven primary levels within the Scottish education system with young people entering primary school at stage one. The school year usually begins in mid-August and due to the admissions criteria based on the month of birth young people usually start school between the ages of 4.5 and 5.5 years old. Parents can request a deferred start date and so some young people can be older. The Growing up in Scotland study into early primary school found that 42 per cent of young people were under 5 when starting school, 49 per cent were aged between 5.0 and 5.5 years, and 9 per cent were older than 5.5 years (Bradshaw et al., 2012).

The age that the learners completed the pre-intervention AfE (InCAS) assessment is shown in Figure 16. Each learner is provided an InCAS login from the University of Durham and the personal details, including age, is populated from the SEEMiS database. 41.8 per cent of the participants were under 7, 48.5 per cent were between 7 and 7.5 and 9.7 per cent were older than 7.5. These ratios are comparable to the data for Scotland's primary schools.

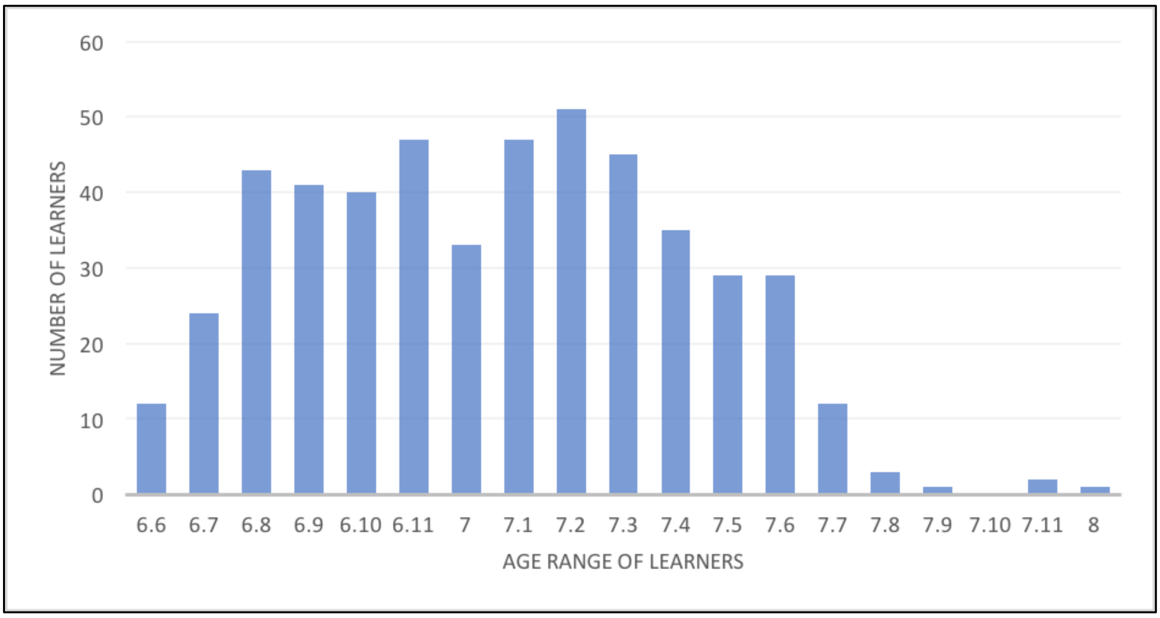


Figure 16: Age range (in years and months) of the learners at the start of the study.

The decimal mean age of the learners at the start of the study was 7.05 with a standard deviation of 0.3. The decimal mean age within the control group was 7.04 with a standard deviation of 0.36 and the intervention was 7.06 with a standard deviation of 0.26.

4.3.6 Attendance

The attendance rates of the learners were identified as a possible extraneous variable. The SEEMiS system was harvested for the attendance records of the individual students at the end of the academic year. The attendance records of the control and intervention classrooms are presented in Figure 17. Although there is a public record held for the attendance figures for each school in Scotland, there is no accumulated data for primary schools. The attendance for the learners in the intervention classrooms was M=94.9, SD=5 and the control group of M=94.5, SD=6.1. The results are slightly higher than the average attendance across all Scottish schools which was 93.7 per cent, but this figure includes secondary schools (Scottish Government, 2015a). The results from a two-tailed independent *t*-test (α=0.05) show that there is not a statistical difference between the attendance rates of the control and intervention classrooms, *t*(493)=-0.774, p=0.439.

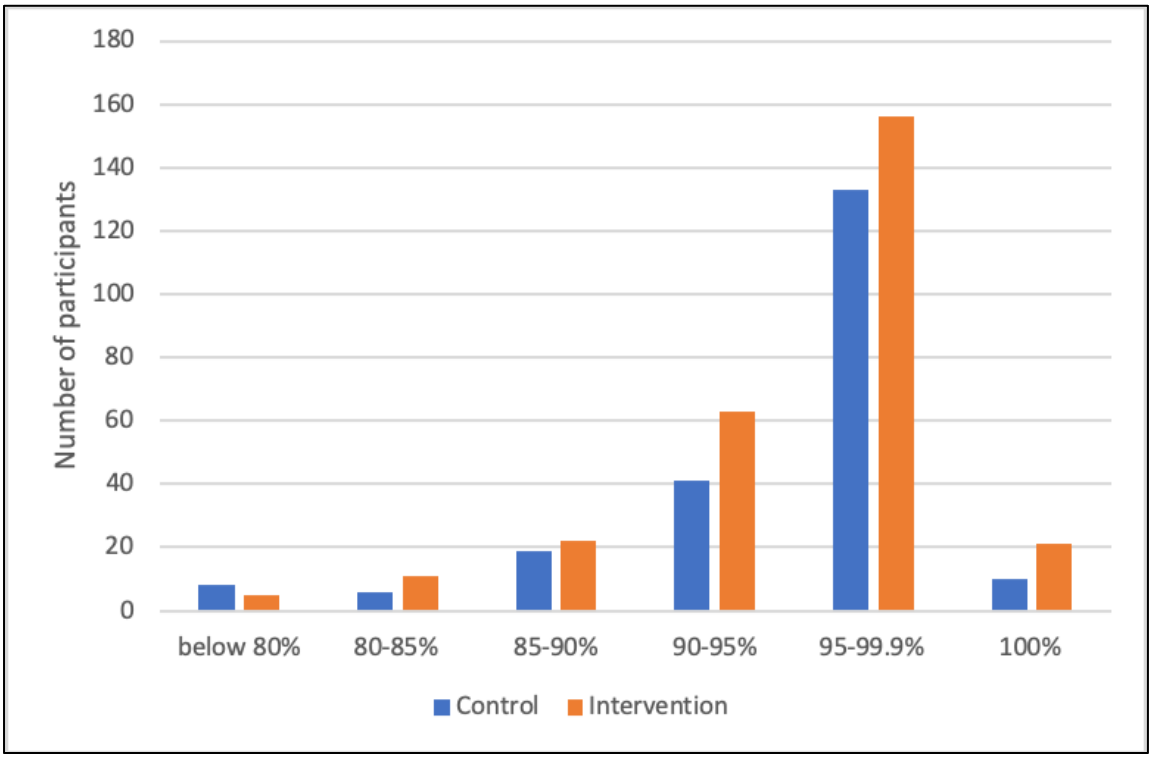


Figure 17: Attendance rates (%) of the learners during the year of the study.

The attendance records for the learners from SIMD 1 in the control classroom were M=92.3, SD=7.3 and M=92.6, SD=5.7 in the intervention classroom. The Scottish Government (2015a) census shows the attendance amongst the 20 per cent most deprived in all Scottish schools was 91.4 per cent. Learners from SIMD 5 had attendance recorded at M=96.9, SD=3.4 in the intervention classroom and M=95.9, SD=5.8 in the control. These are comparable to the national average of 95.7 per cent. Learners from SIMD 5 have a better attendance record than learners from SIMD 1 both nationally and in the research group. Only six learners from SIMD 1 (control=2, intervention=4) achieved 100 per cent attendance compared to fifteen from SIMD 5 (control=6, intervention=9). In contrast, all the learners that recorded attendance levels of below 85 per cent (n=17) came from SIMD 1 (control=8, intervention=9), none were from SIMD 5. The results from a two-tailed independent *t*-test ($\alpha=0.05$) show that there is not a statistical difference between the attendance rates of the control and intervention classrooms based on SIMD quintile: SIMD 1 $t(153)=-0.325$, p=0.746, SIMD 2, $t(74)=1.54$, p=0.127 SIMD 3 $t(24)=-0.18$, p=0.860, SIMD 4 $t(53)=-0.348$, p=0.729 and SIMD 5 $t(181)=-1.52$, p=0.131.

4.3.7 Summary of characteristics of participants

Slack and Draugalis (2001) observes that part of the process of establishing external validity involves examining the characteristics of the participants. The age range, gender, ethnicity and language competency of the 495 learners that participated in this research is comparable to the national data for learners at the primary 3 stage and those who attend all of the Scottish primary schools. The level and type of ASN recorded on the participants is representative of the 57,750 learners in primary 3. The attendance records of the participants are also representative of the whole Scottish school system. There is a higher representation of quintiles 1 and 5 in this research as it was a central component of the study design.

4.4 Characteristics of the Teachers

The registration status, experience, gender and ethnicity of the class teachers in both the control and intervention classrooms was obtained through a questionnaire. The comparable data was based on the Summary of Scottish Statistics from the pupil census and supplementary raw data files from the teacher census 2015 (Scottish Government, 2015a, Scottish Government, 2015c)

Overall, 29 teachers participated in the research, this includes two teachers that left during the intervention period. As Table 8 illustrates both teachers that left were from the intervention classroom. One teacher went on maternity leave six weeks into the intervention and was replaced by the learning support teacher for the remainder of the year. The other went on long-term sick leave twelve weeks into the intervention before leaving her post and being replaced by long-term supply. The two teachers that left their post represents 6.9% of the sample. Data on the characteristics of the teacher that left her post was not collected and this represents a 3.4% attrition rate.

Teacher Attrition Rates			
	All Teachers	***Control***	***Intervention***
Pre-intervention	29	12	17
Teachers left post	2	0	2
Reasons	Maternity leave Resignation following long-term sick leave		
Analysed	28	12	16
Excluded from analysis	1	0	1
Reason	The teacher that resigned following long-term sick leave did not complete the survey.		

Table 8: Attrition rate for the class teachers in the control and intervention classrooms.

4.4.1 Working pattern of teachers

Analysis of the data on the 16 teachers in the intervention classrooms and 12 in the control was undertaken. The control classroom had 12 teachers that worked full-time. The intervention classroom had 11 classes taught by a full-time teacher and two classes taught by two teachers working on a job share basis. The teacher that left on maternity leave worked full-time. The teacher census records that 82% of teachers in the primary system work full-time with 18% part-time (Scottish Government, 2015a). This is comparable to 85.7% and 14.3% of the 28 teachers that were analysed in this research.

4.4.2 Gender ratio of teachers

Eighty-nine point two-nine per cent of the classroom teachers (n=25) were female with 10.71 per cent (n=3) male. The two male teachers were in the intervention classrooms and one to the control. As Figure 18 illustrates this is comparable to the male/female ratio of teachers in the primary sector in Scotland (Scottish Government, 2015c).

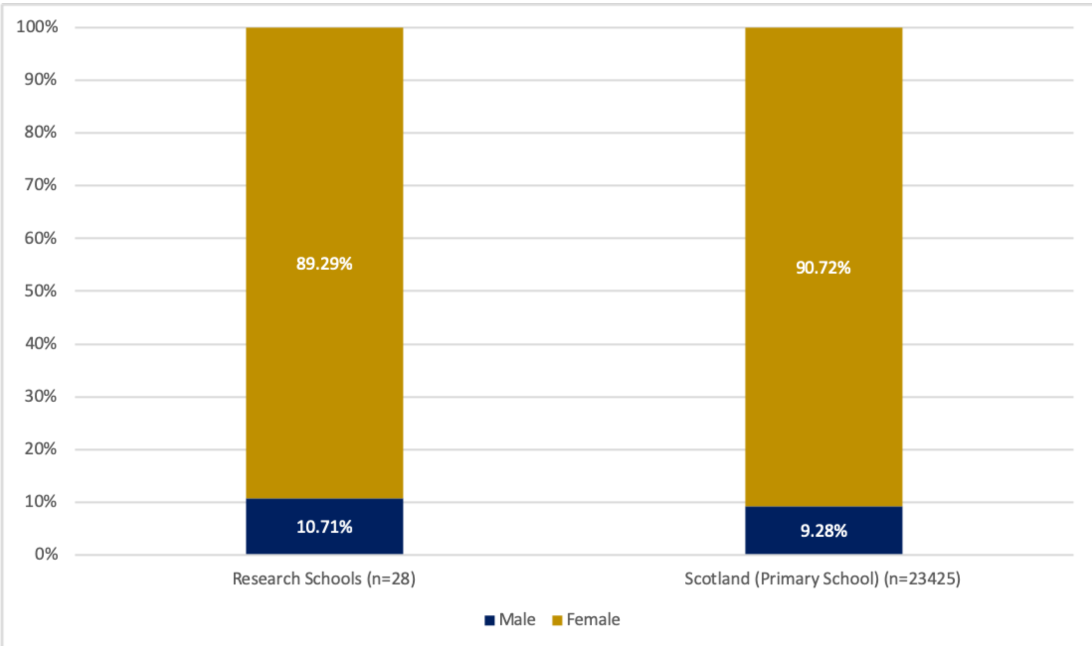


Figure 18: Gender ratio for primary teachers working in public-funded schools based on the 2015 teacher census supplementary data.

4.4.3 Ethnicity of teachers

The teacher questionnaire confirmed that 96 per cent of the primary 3 class teachers in both the control and intervention classrooms were White Scottish or White Other, British. Four per cent was from a minority ethnic group, Asian Other. As Figure 19 shows, these proportions reflect the ethnic distribution of teachers in primary schools in Scotland (Scottish Government, 2015c).

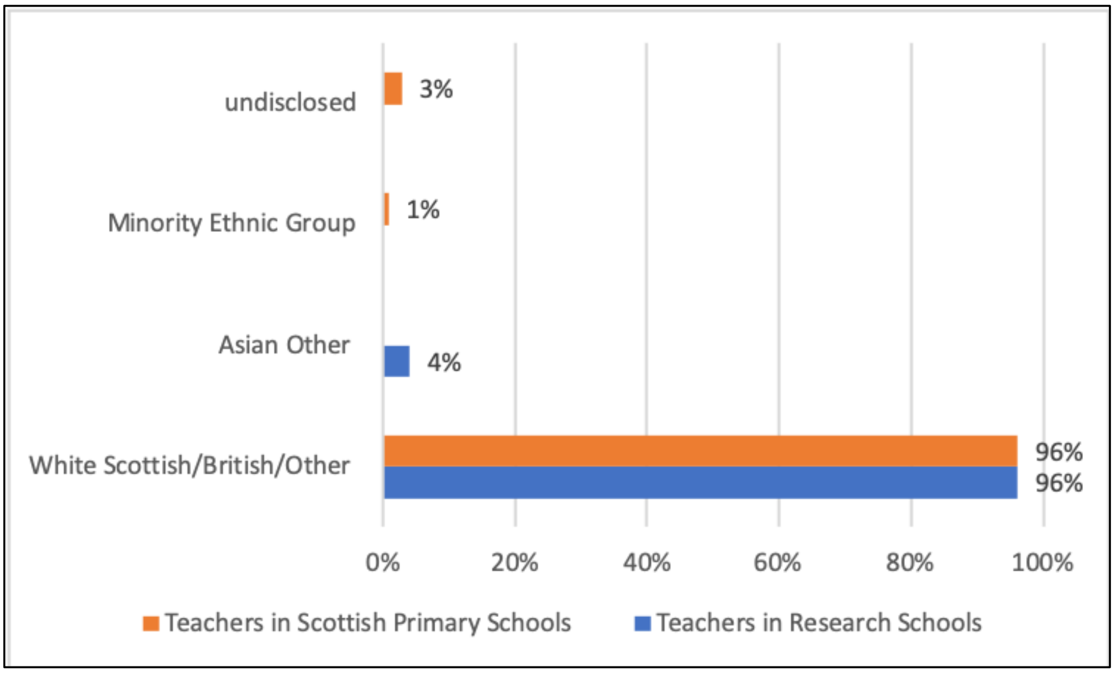


Figure 19: Ethnic distribution of teachers in the research schools compared to Scottish primary schools based on the 2015 Teacher Census.

4.4.4 Registration status and teaching experience

It is a legal requirement for all teachers working in Scottish schools to register with the General Teaching Council of Scotland. A student teacher after completing their initial training obtain provisional registration status and will undertake a probationary period before achieving full registration. Teachers with full registration must maintain their status by engaging in ongoing professional review and development (General Teaching Council of Scotland, 2017).

The teacher questionnaire confirmed that the majority of teachers had full registration status. Three teachers were in their probationary year, two in the intervention classroom and one in the control. As Figure 20 illustrates the range of teaching experience was fairly evenly spread across both the control and intervention groups.

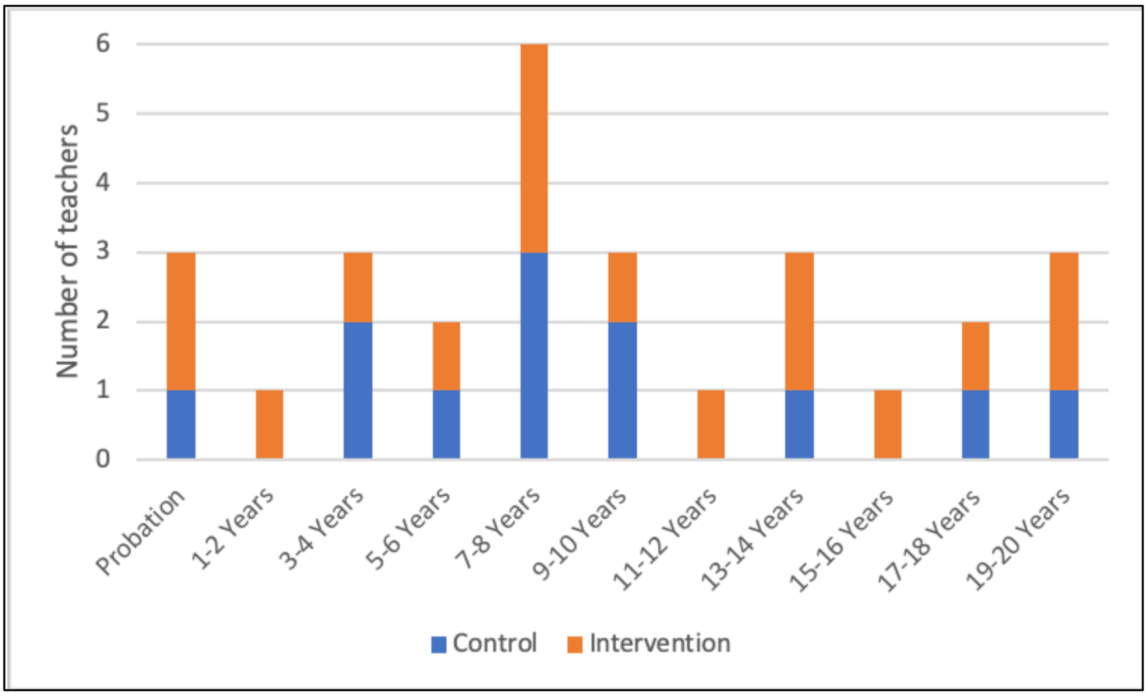


Figure 20: Level of teaching experience of each teacher in the control and intervention classrooms.

4.4.5 Summary of the characteristics of the class teachers

The twenty-eight teachers whose data were analysed in this research are a representative sample of those teaching in primary schools in Scotland. The gender, ethnicity ratio and working pattern of the teachers in the research schools reflect the distribution within the Scottish education system. A broad range of teaching experience is demonstrated across both the control and intervention classrooms.

4.5 Research Schools

As discussed in chapter 3 the schools were recruited using four inclusion criteria: location, SIMD quintile rating, school roll, self-contained classrooms of similar volume/construction that were not adjacent to gyms, halls or lunchrooms.

4.5.1 Location and type of schools

All schools were in the Fife region of Scotland, which is the third largest of thirty-two local authorities and accounts for 6.9% of the total Scottish population (Scottish Government, 2014a). All the research schools were mainstream state-funded. Three schools were Roman Catholic and although they promote the values of the Catholic faith, they are state-funded and follow the national curriculum. Seventeen point nine per cent of all Scottish primary schools are Roman Catholic (Scottish Government, 2015a).

Table 9 shows the four classifiers used to define areas in Scotland. The 2015 pupil census records that 41% of all Scottish schools are in urban areas and 15% in small towns (Scottish Government, 2015a). There is no separate data on primary schools and therefore no target school population to establish external validity. Seventy-seven per cent (N=10) of the research schools were in urban areas and 23% (N=3) were in small towns.

Area classifier	*Population size*
Large urban	125,000 or more
Urban	10,000-124,999
Small towns	3000-9,999
Rural	populations of below 3,000

Table 9: Classifiers used to define areas in Scotland (Scottish Government, 2012a)

4.5.2 School roll and class sizes

As Table 10 illustrates the school rolls ranged from 184 to 398 and this is representative of the school roll in Scotland when small rural schools are excluded. The maximum class size in a single-stage primary 3 class in Scotland is 30 students with a composite class having a maximum size of 25 students (Scottish Government, 2015a). Fife Council applies a capped learner to teacher ratio of 1:18 for primary 1-3 classes in areas of social deprivation. As Table 10 shows the number of participants in the control class ranged from 13 to 28 and the intervention class from 16 to 29. There were thirteen intervention classes (N=278) and twelve control classes (N=217).

Research School	School Roll	Classroom	P3 participants	Condition
School 1	277	1A	17	soundfield
		1B	18	control
School 2		2B	16	soundfield
	235	2A	16	control
School 3	398	3A	29	soundfield
		3B	25	control
School 4	347	4B	28	control
		4C	18	control
		4A	27	soundfield
School 5	393	5B	13	control
		5A	26	soundfield
School 6	208	6A	18	soundfield
School 7	372	7A	27	soundfield
		7B	26	control
School 8	246	8A	23	soundfield
School 9	198	9A	13	control
		9B	16	soundfield
School 10	184	10A	22	soundfield
School 11	221	11A	17	soundfield
		11B	18	control
School 12	339	12A	23	soundfield
		12B	16	control
School 13	273	13A	17	soundfield
		13B	13	control
		13C	13	control

Table 10: School roll and the number of participants in the control and intervention classrooms.

4.6 Room Acoustics

Ambient noise, reverberation times and speech clarity levels were identified as extraneous variables. Comparisons were made to establish if these were evenly spread across the control and intervention classrooms.

4.6.1 Classroom area and ambient noise levels

The same method was used to measure all of the 25 control and intervention classrooms. All measurements were undertaken when the classrooms were unoccupied at the beginning or end of a school day. Each classroom's length, height, and width were measured using a Ridget Micro LM-100 Laser Distance Measure Meter to obtain the room volume (see Appendix 10 for the specification data).

There are no statutory guidelines for room acoustics in Scottish schools although Building Bulletin 1993, which is mandatory in England and Wales is regarded as a key reference document. Building Bulletin 1993 recommends that ambient noise levels are recorded on a dry day, with one person in the room, windows open and the heating system switched on (Department of Education and Skills, 1993). It was not possible to independently control the status of the heating system in the research schools and so all measurements were undertaken in the winter when the heating systems were active. The maximum ambient noise levels recommended for self-contained primary classrooms is 35 dB $L\mathrm{A}_{eq,30min}$ but in rooms with negligible variations in noise levels measuring over a shorter time such as $L\mathrm{A}_{eq,5min}$ provides a reliable indication if the performance standard will be met (Department of Education and Skills, 1993). As none of the schools were subject to intermittent variation in noise levels from airplanes and railways the ambient noise was recorded using $L\mathrm{A}_{eq,5min}$.

Three separate measurements were taken in rooms and these represented positions that would usually be occupied during a lesson. Recording took place at the front, middle and back of the room. The average of these three recordings provided the ambient noise levels for the classrooms and are presented in Table 11. A Casella CEL-620A class 2 sound level meter meeting standards ANSI S1.4 and

IEC 60651 to Type 2 was used to measure the ambient noise levels. In keeping with Building Bulletin 1993 the sound level meter was placed on a tripod at 1.2 meters from the ground and at 1 meter from any furniture (Department of Education and Skills, 1993).

Research School	*Location*	*Class room*	*Condition*	*Volume (m^3)*	*Ambient Noise ($LA_{eq,5min}$)*
School 1	small town	1A	intervention	175	40.1
	small town	1B	control	174	40.2
School 2	urban	2B	intervention	224	48.7
	urban	2A	control	224	48.5
School 3	urban	3A	intervention	215	48.8
	urban	3B	control	215	48.4
School 4	urban	4B	control	282	49.2
	urban	4C	control	206	45.9
	urban	4A	intervention	227	47.7
School 5	small town	5B	control	250	38.9
	small town	5A	intervention	250	39.1
School 6	urban	6A	intervention	200	39.6
School 7	urban	7A	intervention	185	41.1
	urban	7B	control	192	40.7
School 8	urban	8A	intervention	215	43.3
School 9	urban	9A	control	200	41.8
	urban	9B	intervention	169	42.5
School 10	urban	10A	intervention	205	43.5
School 11	small town	11A	intervention	170	43
	small town	11B	control	170	43.2
School 12	urban	12A	intervention	160	48.7
	urban	12B	control	180	43.2
School 13	urban	13A	intervention	188	52.3
	urban	13B	control	188	50.2
	urban	13C	control	188	52.1

Table 11: Level of ambient noise in the unoccupied intervention and control classrooms.

The mean volume for all twenty-five classrooms was V_{mean} =202.08m^3 (SD=29.59). The control classrooms had a V_{mean} =205.75m^3 (SD=33.01) and this compares to the intervention classrooms which had a V_{mean} =198.69m^3 (SD=26.94.) There is no significant difference in the mean values of the classrooms: $t(23)=-0.588$, $p=0.563$. The ambient noise levels in the unoccupied control classrooms ranged from 38.9-52.1 $LA_{eq,5min}$ compared to the unoccupied intervention classrooms of 39.1-52.3 $LA_{eq,5min}$. The mean ambient noise levels for the unoccupied control classroom were 45.2 $LA_{eq,5min}$ (SD=4.4) and this compares to a mean of 44.5 $LA_{eq,5min}$ in the intervention rooms (SD=4.3). The mean ambient noise levels across all conditions were 44.8 $LA_{eq,5min}$ with a standard deviation of 4.2. The results from a two-tailed independent *t*-test ($\alpha=0.05$) show that there was not a significant statistical difference between the mean ambient noise levels in the control and intervention classrooms: $t(23)=-0.403$, $p=0.691$.

4.6.2 Reverberation time and speech clarity

A convolution reverb method was used to obtain the RT_{60} and C_{50} in all the control and intervention classrooms. Apart from the tester, measurements were undertaken when the rooms were furnished but unoccupied, at the beginning or end of a school day. The exponential sine sweep method for impulse measurement was deployed which uses an exponentially swept sinusoid for room excitation, and aperiodic deconvolution to extract the impulse response from the recorded room response (Carson et al., 2009).

4.6.3 Method

The exponential sine sweep was generated and recorded simultaneously using Aurora plugins in conjunction with Audacity 2.0.5 software which was set in overdub mode. The logarithmic sweep had a frequency range of 20-20,000Hz and had a sweep duration of 20 seconds long. The exponential sine sweep recorded waveform was set to mono, with a sample rate of 48 kHz and a resolution of 32-bit float. Figure

21 proves a screenshot from Audacity showing the exponential sine sweep signal, the inverse filter waveform and the recorded waveform of the room.

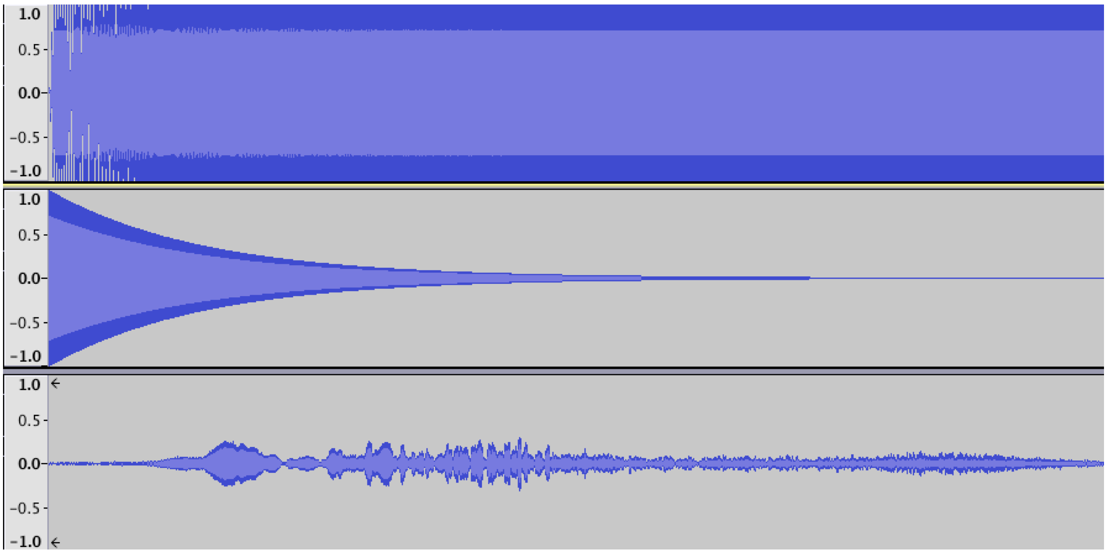


Figure 21: Screenshot from Audacity showing the exponential sine sweep reference waveform (top), inverse filter (middle) and the exponential sine sweep (bottom).

Figure 22 provides a schematic diagram of the method used to obtain the impulse response measurements. The software was installed on a Lenovo T420 laptop running on Windows 7. As the quality of internal sound cards on laptops would introduce an extraneous variable the laptop was connected to a RME Fireface UC USB 2.00 Compact Audio Interface. The exponential sine sweep was emitted through a Behringer Truth B3030A Audio Reference Monitor which was connected via a 6mm audio cable to the monitor output of the RME Fireface UC USB 2.00 Compact Audio Interface. This sat on a speaker stand at a distance of 1 meter from the ground.

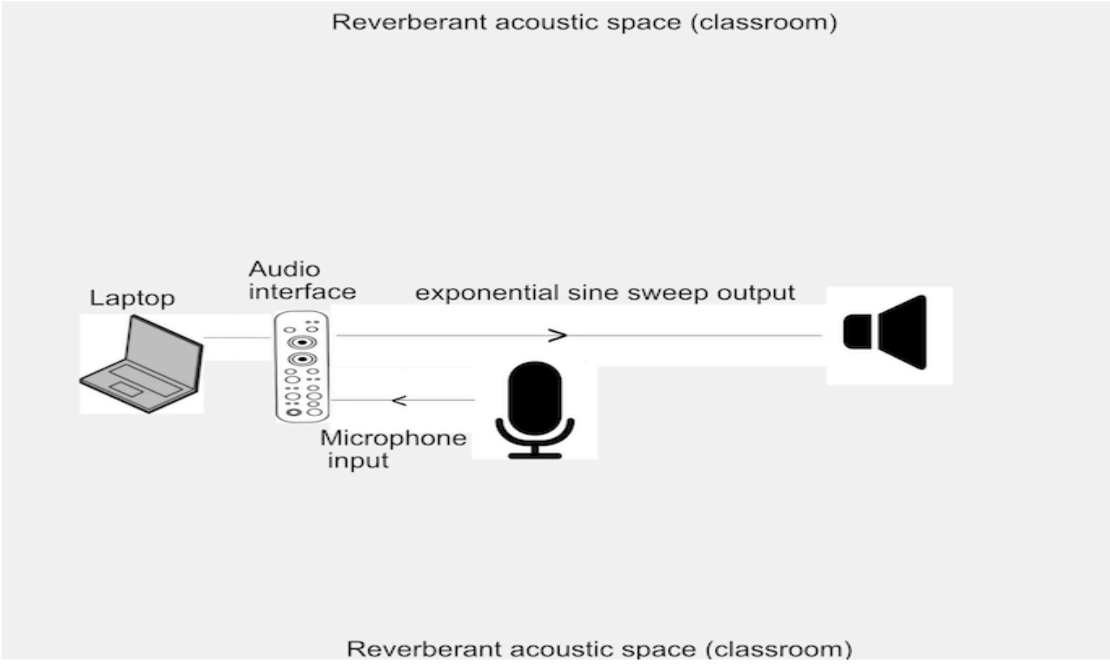


Figure 22: Room set up for recording the RT_{60} and C_{50}.

The excitation of the room was captured by an Audio Technica AT8010 omnidirectional condenser microphone which was plugged into channel 1 of the RME Fireface using an XLR cable (Specification sheets for the speaker, interface and microphone are in Appendix 11). The distance between the microphone and the nearest reflecting surfaces, including the floor was 1 meter. The minimum distance between the microphone and speaker was calculated using ISO 3382 1 (2009) standards, which defines the minimum distance as:

$$d_{min} = 2\sqrt{\frac{V}{cT}}$$

where V is the volume of the classroom c is the approximate value for the speed of sound and T is the estimated RT_{60}. Using the highest room volume ≅ 282m^3, c≅300ms and T=1s, this provides: d_{min} ≅ 1.9m (Quartieri et al., 2009).

$$d_{min} = \sqrt{\frac{282}{300}} = 1.9$$

A distance of 2 meters between the speaker and the microphone was adopted. To obtain a better mapping of the classroom environment two measurements were taken: one in the middle of the classroom and one near the front where the teacher would stand. The values of the two recordings were then averaged per frequency.

The impulse response of each classroom was extracted using the Aurora convolve module in Audacity. The process involves convolving the recorded waveform of each room with the inverse filter function of the exponential sine sweep. As Figure 23 illustrates the impulse response is then edited to remove the sections before the sound was produced and after it had decayed completely. The impulse response is obtained by narrowband filtering the recorded signal across the frequencies. Reverberation times are then obtained for each octave band using the Schroeder backward integration method (Campanini and Farina, 2009).

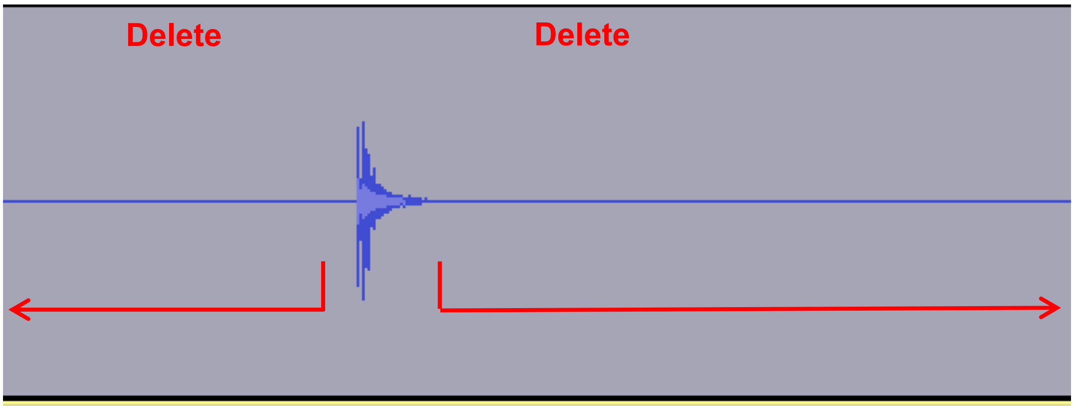


Figure 23: Editing process of the impulse response.

The Aurora plugin provides data on RT_{20} and $RT_{30.}$ Horvat et al. (2007) advises that the RT_{30} parameter is regarded as the most accurate and this should be recorded where possible. The RT_{30} is twice the time needed for the sound pressure level to decrease by 30dB, the rate of decay is measured by the linear least-squares regression of the measured decay curve from - 5 dB to -35 dB (Campanini and Farina, 2009). As it is linear the RT_{60} can be extrapolated from the $RT_{30.}$ The Aurora provides the data in a .txt format and these were imported into Microsoft Excel 2016 for analysis.

4.6.4 Reverberation time results

Building Bulletin 93 provides the performance standards of reverberation for teaching and study places in school buildings using $T_{mf.}$ This is the arithmetical average of the frequencies 500Hz, 1000Hz and 2000Hz. For primary school buildings, the recommended RT_{60} should not exceed 0.6s (Department of Education and Skills, 1993). The mid-frequency RT_{30} along with the room volume for each of the classrooms is shown in Table 12 (a full range of the frequencies for each room are presented in Appendix 12).

Research School	***Classroom***	***Condition***	***Room Volume (m^3)***	***$T_{mf\ (s)}$***
School 1	1A	intervention	175	0.62
	1B	control	174	0.62
School 2	2B	intervention	224	1.03
	2A	control	224	1.01
School 3	3A	intervention	215	0.81
	3B	control	215	0.83
School 4	4B	control	282	0.79
	4C	control	206	0.65
	4A	intervention	227	0.71
School 5	5B	control	250	0.84
	5A	intervention	250	0.87
School 6	6A	intervention	200	0.90
School 7	7A	intervention	185	0.47
	7B	control	192	0.52
School 8	8A	intervention	215	0.75
School 9	9A	control	200	0.80
	9B	intervention	169	0.81
School 10	10A	intervention	205	1.18
School 11	11A	intervention	170	0.69
	11B	control	170	0.68
School 12	12A	intervention	160	0.70
	12B	control	180	0.70
School 13	13A	intervention	188	0.85
	13B	control	188	0.84
	13C	control	188	0.84

Table 12: T_{mf} and room volume of the control and intervention classrooms.

The $T_{mf\ mean}$ for all the classrooms was 0.78s with a standard deviation of 0.16s. The control classrooms had a $T_{mf\ mean}$ of 0.8s and a standard deviation of 0.14s compared to the intervention classrooms where the $T_{mf\ mean}$ 0.76s and a standard deviation of 0.17s.

The findings from this research are in keeping with results from previous studies. Knecht et al. (2002) measured the reverberation times and noise levels in thirty-two classrooms in eight elementary schools in Ohio. The longest reverberation times were found in the classrooms with the highest ceilings and in general, the rooms with the smallest room volume had the shortest reverberation times. The fabric of the building was also a contributory factor with poorly installed single glazed windows associated with longer reverberation times and higher background noise levels. The three classrooms measured in the current study that had reverberation times >1.0s all had ceiling heights in excess of 4 meters. These rooms also had large single-pane windows. Although no detailed survey of the school building was undertaken the longest reverberation times were in older school buildings, conversely the newer build schools had shorter reverberation times.

Klatte et al. (2010b) observe the acoustic conditions under which learning takes place influences the learning process. Long RT can reduce speech intelligibility and so could be an extraneous variable that could compromise validity if not evenly distributed across the control and intervention classrooms. Figure 24 provides the T_{mf} displayed in a stacked bar chart format to illustrate the even spread of the different acoustic conditions across the control and intervention classrooms.

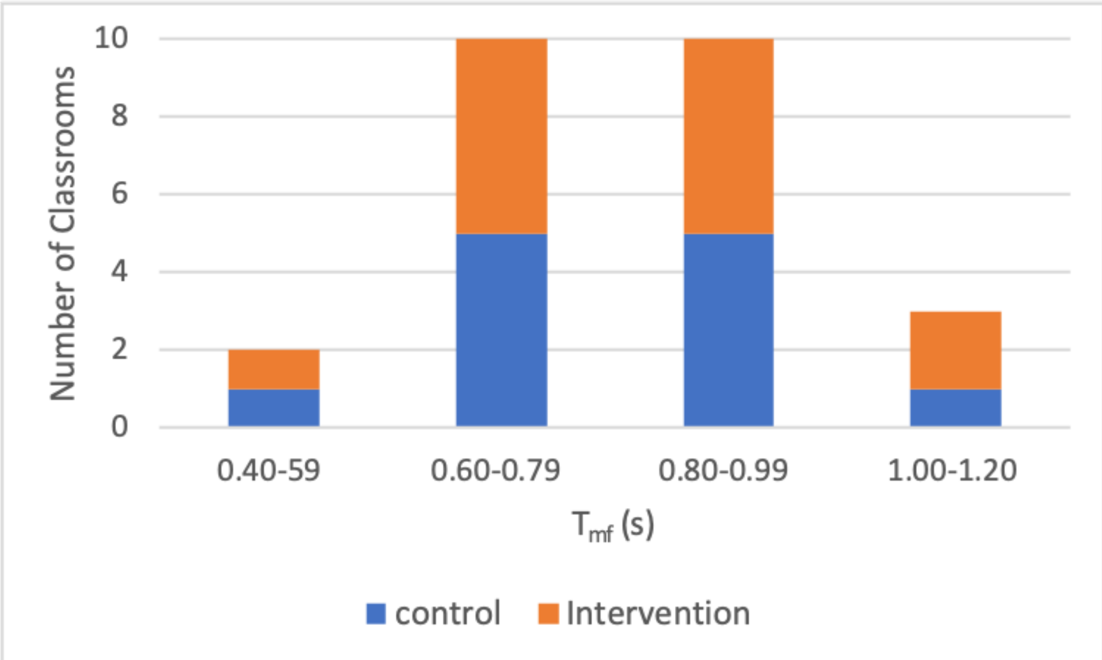


Figure 24: Distribution of T_{mf} across both the control (blue) and intervention (orange) classrooms.

Only two classrooms had short reverberation times of <0.6s which comply with the standard for primary schools and these are spread across both conditions. Three classrooms had RT above 1.0s, two had dynamic soundfield fitted and the other was a control room. The majority of the classrooms (N=20) were between 0.6s and 0.9s and this was fairly evenly spread across both conditions. One criticism of the T_{mf} method for measuring a room for speech intelligibility is that it does not always correlate with the subjective evaluation of classrooms by the listener. A study of 110 classrooms in Warsaw evaluated the effects of reverberation time on speech, using the speech transmission index found that T_{wf} provided a more accurate assessment of the acoustic properties of a classroom than $T_{mf.}$ T_{wf} is calculated in the same manner as T_{mf} but using a wider range of frequencies (250Hz, 500Hz, 1000Hz, 2000Hz, 4000Hz) (Mikulski and Radosz, 2011). Interestingly, this is the same frequency range that the British Society of Audiology uses to define hearing thresholds (BSA Professional Practice Committee, 2015).

Table 13 shows the difference between mid-frequency and wide-frequency parameters derived from the RT_{30} impulse response recorded in each control and intervention classroom. The $T_{wf\ mean}$ for all the classrooms was 0.72s with a standard deviation of 0.11s. The control classrooms had a $T_{wf\ mean}$ of 0.73s and a standard deviation of 0.13s compared to the intervention classrooms where the $T_{wf\ mean}$ 0.72s and a standard deviation of 0.11s. The change in scores from T_{mf} to T_{wf} is marginal with the mean difference between the intervention and control rooms recorded as 0.09s and 0.06s. The standard deviation for both is 0.04s. As expected, the biggest difference occurs in the rooms with longer reverberation times.

Research Schools	***T_{mf} (s)***	***T_{wf} (s)***	***Difference (s)***
school 1 intervention	0.62	0.65	0.03
school 1 control	0.62	0.65	0.03
school 2 intervention	1.03	0.94	0.09
school 2 control	1.01	0.90	0.11
school 3 intervention	0.81	0.72	0.09
school 3 control	0.83	0.77	0.06
school 4 intervention	0.71	0.68	0.03
school 4 control	0.65	0.65	0.00
school 4 control	0.79	0.76	0.03
school 5 intervention	0.87	0.81	0.06
school 5 control	0.84	0.77	0.07
school 6 intervention	0.90	0.78	0.12
school 7 intervention	0.47	0.49	0.02
school 7 control	0.52	0.59	0.07
school 8 intervention	0.75	0.69	0.06
school 9 intervention	0.81	0.77	0.04
school 9 control	0.80	0.77	0.03
school 10 intervention	1.18	1.01	0.17
school 11 intervention	0.69	0.66	0.03
school 11 control	0.68	0.60	0.08
school 12 intervention	0.70	0.61	0.09
school 12 control	0.70	0.61	0.09
school 13 intervention	0.85	0.70	0.15
school 13 control	0.84	0.71	0.13
school 13 control	0.84	0.71	0.13

Table 13: Difference between the T_{mf} and T_{wf} parameters for evaluating the performance standards in school classrooms.

Figure 25 provides the T_{wf} presented in a histogram to illustrates that a similar even spread across the different acoustic conditions is maintained. The main difference when using T_{wf} instead of the T_{mf} standard is that only one intervention classroom has a reverberation time of >1.0s instead of three. There is an increased number of schools (N=22) that range from 0.6s and 0.9s with a higher percentage between 0.6s and 0.7s.

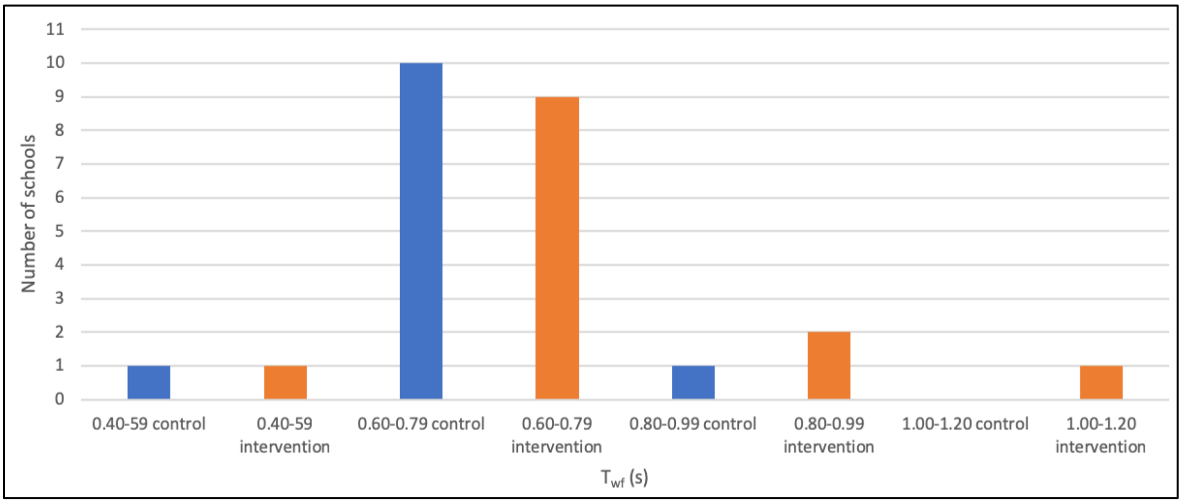


Figure 25: Distribution of T_{wf} reverberation times across both the control (blue) and intervention (orange) classrooms.

4.6.5 Weighted speech clarity and speech intelligibility

Although reverberation times are a predominant indicator for describing the acoustic properties of a room there are a number of studies that suggest that the early/long energy ratios may better specify the effects that a room has on speech quality and intelligibility (Haas, 1972, Marshall, 1995, Bradley et al., 1999, Bradley et al., 2003). Haas (1972) investigated the effects of early arriving reflections on speech and found that to the listener early acoustic reflections do not appear as a separate sound source but instead integrate with the direct sound and so enhance the loudness and sound quality of speech, conversely, longer reflections appear separate and are detrimental to speech intelligibility. Previous studies have highlighted the contribution of early reflected sounds on improved SNR. Bradley et al. (2003) examined the effects of increased early reflected speech sound energy on speech intelligibility using subjective speech assessments and impulse response measurements in a simulated room. The results indicate that early reflections provide an effective SNR benefit of 9dB and in situations where the direct sound is compromised, they are essential to successfully access spoken communication.

The participants in the study were both hearing and deaf with a mean age of 28 and 60 years respectively which is a higher age range than the current cohort, however as young learner's auditory system is subject to development it can be assumed that younger people would also receive a SNR advantage similar or better. Therefore, a metric that provides the statistical distribution of the effects of early reflected sounds across the control and interventions classrooms was factored into the study design.

C_{50} is an objective parameter that compares the ratio of early to long reflected sounds in an impulse response to determine the effects the acoustic properties of a classroom are having on speech clarity. C_{50} measures reflections that arrive within 50 milliseconds of the direct speech sound and are mainly perceived as if the direct sound source has been amplified and enhanced relative to long reflections and so provides a reasonable descriptor of speech clarity (Gade, 2014). Clarity is measured using a logarithmic scale that ranges from negative to positive, a higher value of C_{50} is indicative of the prominence of early reflections and represents a higher level of speech clarity. C_{50} values of 3dB or more are considered favourable to speech (Bradley et al., 2003, Roy and Browne, 2010).

One of the limitations of the C_{50} parameter is that there is no clearly defined method to combine the different octave bands into a single frequency average value for classrooms. ISO 3382 1 (2009) recommends that both 500Hz and 1000Hz represent the mid-frequency range for performance areas but this is a narrow definition for speech in classrooms. Several researchers, most notably Marshall (1995) have advocated applying speech-weighting to the C_{50} values which allow the early/long energy to be defined as useful/harmful ratios, which then can predict both clarity and intelligibility using a rating scale. The weighted C_{50} correlates with other objective measures such as the Speech Transmission Index (STI). The STI evaluates intelligibility based on how well preserved the amplitude modulation of a signal is between teacher to learner, learner to teacher and learner to learner

(Department of Education and Skills, 1993). It is rated 0-1 with 0 being defined as bad and 1 as excellent. Marshall (1995) categorised the speech-weighted C_{50} against the STI which ranged from bad (-7 to -12) to excellent (+7-+18). However, as it is unlikely that any classroom for speech would have such a negative classification, a modified version has been adopted, illustrated in Figure 26 below. This scale was used to classify the speech clarity and intelligibility of the control and intervention classrooms.

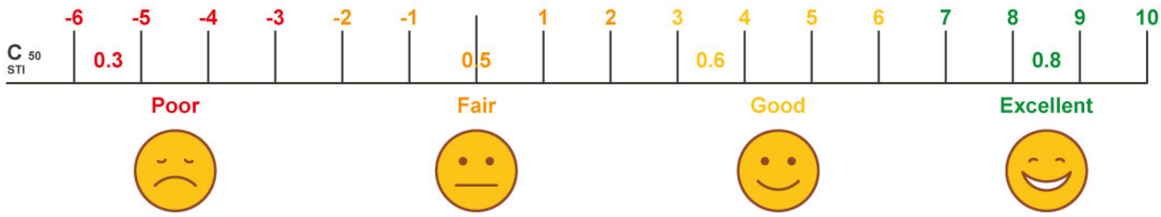


Figure 26: A rating scale for predicting the level of clarity and STI in a classroom using speech-weighted C_{50} values.

4.6.6 Weighted C_{50} analysis and results

In addition to RT_{30}, the Aurora suite of plugins for Audacity 2.0.5 software has a range of parameters suitable for assessing the acoustic properties of a room, these include C_{50}. The data from the impulse response generated using the convolution reverb method described in section 4.6.3 was used to obtain the C_{50} data for the control and intervention classrooms. The C_{50} data is provided across seven octave bands: 250Hz, 500Hz, 1kHz, 2kHz, 4kHz, 8kHz, and 16kHz. The ratio of early to long reflected sounds quantifies the classroom's acoustic properties relating to clarity and reverberation. Clarity is to a large extent indicative of the balance between early reflections and reverberation and an ever-present variable is frequency. The T_{mf} range used to determine RT_{30} is the arithmetical average of the frequencies 500Hz, 1000Hz and 2000Hz and so provides limited information on high frequencies sounds commonly associated with consonants and the second formants in some vowels. Marshall (1995) defined speech values from the energy-time data in four different octave bands: 500Hz, 1, 2 and 4kHz. To best correlate C_{50} with the Speech Transmission Index the octave bands 500Hz, 1, 2 and 4kHz are weighted by 15%, 25%, 35%, and 25% respectively then summed to obtain a composite score. The summing of the C values (in dB) across these four octave bands are then divided by four. The Aurora plugins stores the data in a .txt format and these were imported into Microsoft Excel 2016 where the weighted C_{50} was calculated.

Table 14 shows the C_{50} properties of the classrooms in relation to the T_{mf} and room volume (Appendix 13 provides the C_{50} values for each classroom across the four different frequencies). The results show that the highest C_{50} recordings (school 7) were in the classrooms with the lower T_{mf}. Conversely, the lowest C_{50} recordings (schools 2 and 10) had T_{mf} >1s. The room volume for each of these classrooms were >200m^3.

Research School	Condition	Room Volume (m^3)	$T_{mf\ (s)}$	Weighted $C_{50\ (dB)}$
School 1	intervention	175	0.62	6.94
	control	174	0.62	6.44
School 2	intervention	224	1.03	2.15
	control	224	1.01	2.12
School 3	intervention	215	0.81	3.75
	control	215	0.83	4.22
School 4	control	282	0.79	4.07
	control	206	0.65	5.96
	intervention	227	0.71	4.93
School 5	control	250	0.84	3.97
	intervention	250	0.87	3.92
School 6	intervention	200	0.90	4.63
School 7	intervention	185	0.47	7.53
	control	192	0.52	7.21
School 8	intervention	215	0.75	4.21
School 9	control	200	0.80	3.44
	intervention	169	0.81	4.16
School 10	intervention	205	1.18	1.14
School 11	intervention	170	0.69	2.76
	control	170	0.68	2.85
School 12	intervention	160	0.70	4.39
	control	180	0.70	6.47
School 13	intervention	188	0.85	4.03
	control	188	0.84	4.37
	control	188	0.84	4.73

Table 14: Speech weighted C_{50} properties of the classrooms.

Figure 27 provides the distribution of speech weighted C_{50} values in a histogram to illustrate the even spread of early reflections across the control and intervention classrooms. Using the rating scale in Figure 26 the results indicate that the majority of classrooms (N=18) are rated good for early reflected sounds and that this is equally distributed across both conditions. Only two classrooms are rated as excellent and again this is evenly distributed across both conditions. Five classrooms are rated as fair, three in the intervention and two in the control. Overall, the data indicates that there is an even spread of speech weighted C_{50} and RT_{30} across both the control and intervention classrooms.

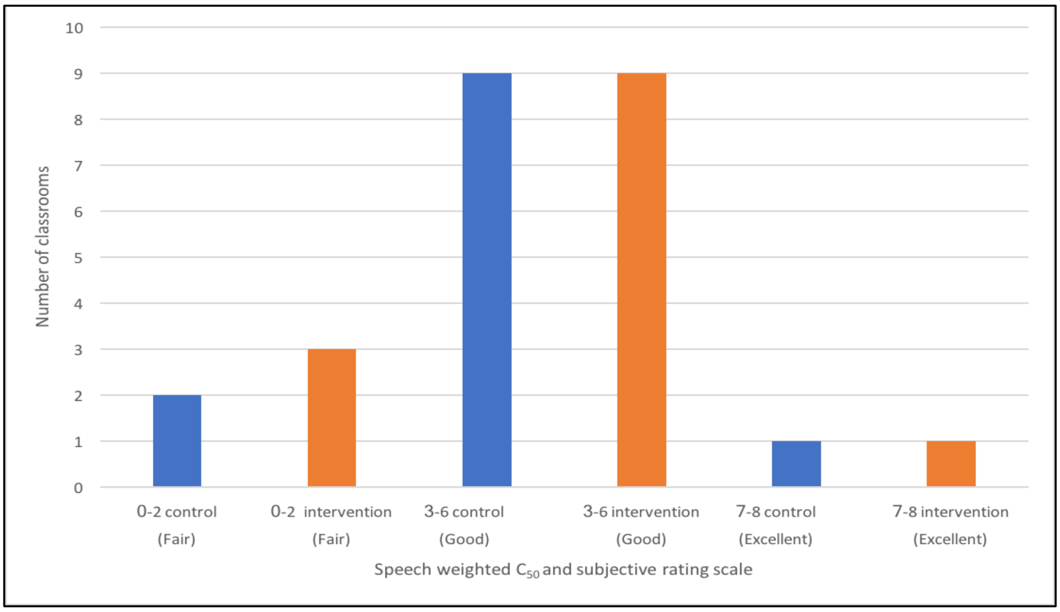


Figure 27: Distribution of speech weighted C_{50} values across the control (blue) and intervention (orange) classrooms.

4.7 Dynamic soundfield Use

In order to gain an insight into the use of the dynamic soundfield systems in the intervention classrooms data was collected on the percentage of days the system was used during the study and the proportion of time each day the device was worn by the teacher compared to the amount of time, the class was in their classroom. The next sections detail the data collection, analysis, and results.

4.7.1 Daily use data collection

Two methods were used to establish net use of the system: Datalogging and weekly timetables. The Inspiro transmitter of the Digimaster 5000 soundfield has an integrated Datalogging system and the data from this can be downloaded and stored through the manufacturer's software. All data was collected and stored on a Lenovo T420 laptop running on Windows 7 which had FM Successware software (version 4.6.3) installed. Once the transmitter data is downloaded the software provides information on how often the system had been used during the day in 15-minute timeframes; the noise levels in the classroom and what audio inputs have been active. The Inspiro transmitter has an inbuilt clock and calendar and so data can be viewed with reference to specific dates and times as well as an overview of a 7-day and 31-day period. The information on the transmitter can only be stored for a 31-day period and so as part of the study design, each system had the data collected, downloaded and stored every 30 days. At this visit, the clocks on the Inspiro were checked against the time on an Apple iPhone 4s and adjustments made, as appropriate.

Soundfield is technology usually installed in a classroom but during the day a class can often be either in other parts of the school or out with the school building. To establish the net use of the systems the study design incorporated the collection of the teachers' timetables. Each teacher in the intervention classrooms provided a master timetable at the start of each term. As primary timetables can be subject to change, in addition each class teacher provided a weekly update. They either gave copies of their weekly timetable sheet or personal planner (n=10) or emailed a list of changes every week (n=2) or sent changes when these occurred (n=3).

The data on the timetable was used in two ways: firstly, when cross-referenced with the Datalogging information it indicated the amount of time during the school day when the class were not in the classroom (physical education, drama, assembly) and this established the net use of the system each day. Secondly, it provided details of when the class was not in the room for a full day (school trips and events) and so could not use the system. This allowed a distinction to be made between teachers not using the dynamic soundfield and not having access to it.

4.7.2 Data analysis

The FM Successware software (version 4.6.3) provides accumulated daily use in a chart or table format. Figure 28 is a screenshot from one school's weekly Datalogging showing the various data formats for illustrative purposes. The histogram chart provides accumulated use, the noise table provides data in 15-minute timeframes colour coded to indicate if the noise levels are above/below 65dB SPL and the list table format provides information on when the system was turned on and off. The charting method was not used as some teachers forgot to turn the transmitter off when charging the system or left it on overnight and so the daily use in this format was higher than that of the school day. To obtain an accurate record on the use of the dynamic soundfield the noise monitoring table was cross-referenced against the listed table and the fifteen-minute timeframes were manually accumulated from the start and end of the school day.

The data from each day's use in the individual school was then inputted into Microsoft Excel 2016. The school day in Fife is five hours and this represented the maximum possible time that the soundfield system could be used. The amount of time that the class was not in the room each day was harvested from the weekly timetable and this was recorded along with the reason (Physical Education, Drama, Assembly, ICT or school trips). Formulas within Microsoft Excel specific to time calculations and percentages were then used to calculate the net daily use of the system.

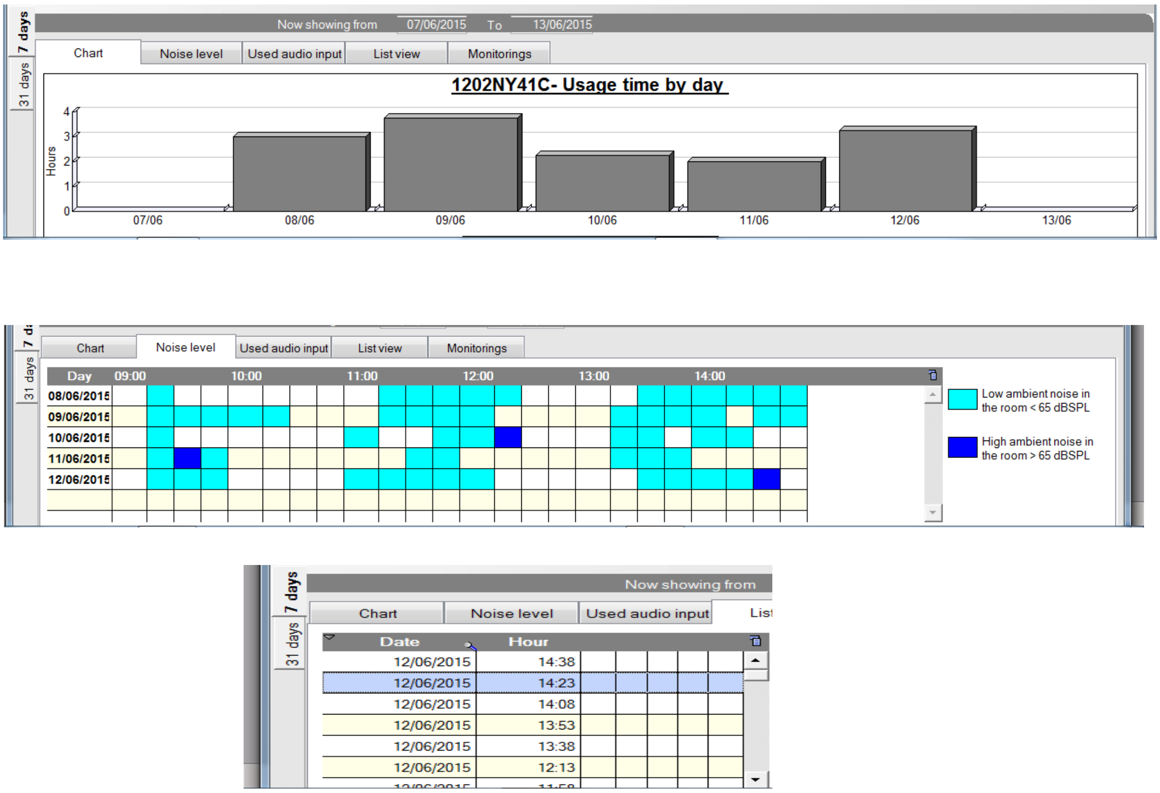


Figure 28: Screenshot from the FM Successware version 4.6.3 showing daily use in a histogram format (top), noise monitoring levels of low and high (middle) and a list table showing when the transmitter was active (bottom).

4.7.3 Daily use results

The dynamic soundfield systems were installed in the classrooms at the end of the pre-intervention stage. The pre-intervention stage lasted twenty-five school days and involved obtaining the learners' consent, hearing screening and the completion of the AfE (InCAS) baseline assessments. Excluding school holidays and teacher training days, the dynamic soundfields were installed until the end of the assessment period and this totalled 154 school days. Young people in Scotland attend school for 190 days a year and so the study lasted 94.2 per cent of the academic year (179 days) with the soundfields installed for 81 per cent of that time. To determine the use of the dynamic soundfield systems in the thirteen intervention classrooms quantitative data was analysed on the total number of days the system was used and the average duration of use as a proportion of the time the class was in their room.

A review of the weekly amendments to the master timetables provided by the schools highlighted a number of days throughout the year when the class was not in their room for a full day. The main reasons were trips, school events and outdoor learning days. The most disruptive event to the master timetable was Christmas with activities such as show rehearsals, singing, fayres and parties resulting in the class being absent from their room for part or all of the day. Although some of these activities lasted half-a-day when this occurred on a day when there was Physical Education or another curricular activity that was outside the classroom the outcome was that the dynamic soundfield could not be used. Furthermore, as you would expect when using technology over a significant period of time faults would arise. As part of the study design, a stock of spares was available and the longest a class was without a system was two days. The two most common faults were charging of the internal battery pack and teachers accidentally cutting the microphone cord during lessons. Not all the reported faults were due to the equipment being broken as in some cases the transmitter required only to be re-connected to the speaker or the volume control had been inadvertently adjusted. As these were perceived and reported as faults and consequently resulted in support these were collectively classified as faults. When calculating the total number of days the system was used a distinction was made between the days when the teacher could not use the system as the class was timetabled elsewhere or there was a reported fault and days it should have been used but was not.

Table 15 provides a detailed summary of the overall number of days the dynamic soundfield system was in use by each intervention classroom during the longitudinal study. The results illustrate that no school dropped out of the study during the academic year. All except two schools achieved daily use of the system at 90% or more with three schools recording 99% and a further three achieving 100% use. Two schools had use of between 86-88%. The analysis showed that one of the schools had teachers that worked on a job share basis and the other had a teacher that left after being on long-term sick leave, the school had irregular supply cover for a period of time. One school was in SIMD 5 area and one in SIMD 1. Where reasons were provided on the amended timetables, teacher absence was the primary cause for the system not being used. Excluding the days where there were reported faults and the class were not in their room, the soundfield was available for a total of 1931 days across the thirteen schools. Table 15 documents that the soundfield was not used for a total of 95 days which represents 4.9%.

	Installed (days)	*Faults (days)*	*Timetabled out of class (days)*	*Reason*	*Available (days)*	*Not used*	*Reason*	*Used (days) (%)*
School 1	154	1	2	1,2	151	0		151 100%
School 2	154	0	5	1,2,3	149	1	A	148 99%
School 3	154	0	1	2	153	0		153 100%
School 4	154	4	6	1,2,3,4	144	14	A	130 90%
School 5	154	2	6	1,2,6	146	21	A,B	125 86%
School 6	154	1	3	1,2	150	0		150 100%
School 7	154	4	5	1,2	145	15	A,B	130 90%
School 8	154	0	4	1,2,6	150	2	B	148 99%
School 9	154	2	5	1,2	147	11	A,B	136 92%
School 10	154	0	2	1,2	152	2	A	150 99%
School 11	154	2	6	1,2,3,7	146	18	A	128 88%
School 12	154	0	4	1,2,5	150	5	B	145 97%
School 13	154	0	6	1,2,3,4	148	6	A,B	142 96%

1=school trips, 2=Christmas activities, 3=sports days, 4=health week, 5=outdoor learning, 6=Easter events, 7=class painted, A=no reason provided, B=teacher absent

Table 15: Net number of days each class were in their room and the net number of days each dynamic soundfield device was used.

To calculate the amount of time the system was used each day in the classroom the total number of minutes of the school day was arithmetically subtracted from the minutes the class was timetabled outside the room. The master timetables and the weekly amendments provided the time of day, duration and reason. School assemblies and PE were a common reason for learners not to be in the classroom across all schools with drama, music, and ICT also being recorded. On days the system was not used without a reason a 0-time duration was recorded. The daily use was inputted into Excel 2016 and the accumulated annual data was analysed. Table 16 provides a breakdown of the average use of the soundfield system during the study along with the total hours the system was used.

	Soundfield available (days)	*Average use (%)*	*Total use (hours)*	*Mean*
School 1	151	67.8	375:25	149.1
School 2	149	69.9	409:20	164.8
School 3	153	63.9	433:20	169.9
School 4	144	75.2	444:50	185.4
School 5	146	53.5	255:10	104.9
School 6	150	80.7	473:10	189.3
School 7	145	70.9	338:45	140.1
School 8	150	71.7	417.00	166.8
School 9	147	60.5	345.00	140.8
School 10	152	64.3	359:30	141.9
School 11	146	60.4	347:35	142.8
School 12	150	93.1	562:30	225.1
School 13	148	73.8	410:25	166.4

Table 16: Average use of the dynamic soundfield system during the longitudinal study.

In considering the average use of the devices a possible trend emerges. Two schools have a particularly high and low average use, and this appears to be influenced by the number of days the system was not used without justification. The teacher in school 12 used the system every day and it was only not used for five days when she was away on a residential trip with the primary 7 class. This compares to school 5 where the teacher went on long-term sick leave and then finally left the school. Short-term cover and a delay in appointing a permanent supply teacher contributed to inconsistent use. Furthermore, the 21 days when the system was not used reduced the overall average score.

In total eight schools (62% of the total research schools) achieved an average use in excess of two-thirds of the school day. The findings from this study are comparable or better to the results from previous studies. Heeney (2007) recorded that 63% of the teachers used the soundfield consistently however, this data was collected from a questionnaire that offered three options: consistently for most sessions, consistently for selected sessions and inconsistently. Not only were the teachers not blind to the intervention and so the process was subject to response and expectation bias, but two-thirds of the scenarios presented to the teachers were positive. Furthermore, there was no definition of the term consistent.

Table 16 also illustrates that four schools (31% of the total research schools) used the system between 60.4%-64.3% of the day. For schools 9 and 11, the most significant factor in lowering the average use scores was the high number of days the system was not used. In addition, an analysis of the Datalogging data indicates that some of the teachers were turning the system off for short periods of time during a lesson. This may have been when the direct teaching component was completed. Dockrell and Shield (2012) also used a questionnaire to establish the usage of the soundfield system in their research with the eleven teachers involved reporting that they used the system for at least 40% of the day. Similar to the Heeney (2007) research the teachers were not blind to the intervention and so the answers may be subject to response and expectation bias. All the schools in this research achieved an average use in excess of 50%.

4.8 Key findings

- The gender and ethnicity characteristics of the learners and teachers in this research are representative of the primary 3 stage in Scotland and the whole of the primary school sector.

- The language levels of learners who have English as an additional language are representative of the whole Scottish primary school sector.
- The percentage of learners with an additional support need is representative of young primary learners in Scotland.
- The attendance figures are representative of the Scottish school population. Learners from the least deprived areas were more likely to achieve full attendance and those from the most deprived had the lowest attendance figures. This was also representative of the Scottish school population. There was no significant difference between the mean attendance rates of the control and intervention groups.
- The characteristics and working patterns of the teachers are representative of the Scottish primary sector.
- There is an even distribution of teaching experience in the control and intervention classrooms.
- The T_{mf} and speech weighted C_{50} values are fairly evenly distributed across both control and intervention classrooms.
- The majority of the classrooms were rated as good for speech weighted C_{50}.
- This research study was unique in the level of detail and data gathered on the use of the dynamic soundfield both through the Datalogging system and class timetables. The results demonstrate no school dropped out of the research during the academic year.
- The dynamic soundfields were available to the classes on 1931 days and only 95 days (4.9%) was the system not used.
- The majority of the schools achieved a level of use in excess of two-thirds of the school day.

The next chapter will analyse the noise surveys and questionnaires undertaken in the control and intervention classrooms.

Chapter 5
Results – Noise Surveys and Questionnaires

5.1 Aims of the chapter

This chapter will address secondary research aims one and four of this thesis (see section 3.2.1). The overarching aim of this chapter is to better understand the noise levels in the control and intervention classrooms. Thereafter, to explore the views and experiences of teachers and learners on the impact of the dynamic soundfield technology when listening in noise. Chapter 4 presented data on the acoustic characteristics of the twenty-five research classrooms. This chapter reports the procedures, analysis, and results from the three instruments used to establish an insight into the level of noise in classrooms: noise surveys of occupied classrooms, learner and teacher questionnaires.

Section 5.2 will present the results from the 150 noise surveys completed in the occupied research classrooms. The purpose was to determine if there was a significant difference in the noise levels during lessons in numeracy, literacy, and IDL in the control and intervention classrooms. In addition, the aim was to establish the level of noise during different learning activities. Section 5.3 provides a description of the purpose, format, data collection and results from the 495 noise questionnaires completed by the research participants. The purpose of the learner questionnaire was to identify the sources of classroom noise, perception of noise levels during lessons and attitudes towards the noise to establish if there was a significant difference between those exposed to dynamic soundfield technology and the control group. Section 5.4 will follow a similar format when presenting the results from the noise questionnaires completed by the 28 class teachers. Section 5.5 will present the results from both the teachers' and learners' questionnaires on the effects of dynamic soundfield on listening in the classroom. The purpose was to evaluate if dynamic soundfield technology improved listening to speech in noise. Throughout, the results will be presented and considered with reference to the relevant literature.

5.2 Classroom noise surveys

5.2.1 Data collection method

Classroom noise surveys were undertaken in all twenty-five control and intervention classrooms. All classrooms were furnished and occupied. Three curricular areas were surveyed: numeracy, literacy, and IDL. These were selected as they are the core subjects that are tested, either directly or indirectly in the AfE (InCAS) assessments which were administered at the pre and post-intervention stage. A Casella 620A Integrated Digital Sound Level Meter (Class 2) which conforms to standards ANSI S1.4, ANSI S1.43, IEC 61672, IEC 60651 and IEC 60804 was used, and this was fitted to a tripod, one meter from any reflective surface at the back of each classroom. Before and after each visit the sound level meter was run through the Casella 120/2 Acoustic Calibrator which conforms to EN (IEC) 60942: 2003, ANSI S1.40: 2006 (Class 2) standards to ensure that consistent and accurate readings were obtained. There was no significant drift in calibration during the assessment period. Measurements were captured at two-minute intervals throughout each lesson as previous research has indicated that such a timeframe provides a good indication of noise fluctuations during the school day. The surveys were completed in both the morning and afternoon sessions as it has been shown that there is no significant difference in noise levels during different parts of the school day. Furthermore, variation in noise levels was approximately constant and so the data collected was confined to the L_{Aeq} and L_{Amin} levels (Shield and Dockrell, 2004).

To ensure a consistent record was achieved each of the three curricular areas was surveyed twice in all the control and intervention classrooms. A total of six surveys per classroom with an overall total of 150 performed. Each activity within a lesson was allocated a unique code, which was modified based on the work of Shield and Dockrell (2004). In their study, six activities that are commonly observed during a lesson were identified. Group work was allocated two categories, a distinction being made between sedentary group work, with some talking and group work with movement and talking. In this research, such differentiation was not made as most group work involves a degree of movement or activity. Furthermore, Shield and Dockrell (2004) did not distinguish between directly supervised and unsupervised individual learning activities that involved some movement. During lessons, teachers can often be working with an individual or group of learners whilst the rest of the class work independently. As the class is not directly supervised and the teacher is focused on other work this can result in more class movement and noise. Table 17 provides a breakdown of the learning activity codes used to classify classroom activities in this research:

Activity Codes	***Activity***	***Description***
A1	Test	Learner working individually, quietly at their table e.g. reading or test.
A2	Individual - One person talking	Learners sitting at table/floor, one person talking (Teacher or Learner) e.g. didactic teaching.
A3	Individual - More than one person talking	Learners sitting at tables/floors, some discussion/talking e.g. class have been set a task/teacher talking to class – more than one person talking.
A4	Individual - More than one person talking; movement	Learners working individually with some talking and movement e.g. class set task, but some individuals are moving around the room.
A5	Individual - More than one person talking; Movement; teacher working with another group/individual	Teacher working with a group or individuals, the rest of the class working at tables. individually, some talking and movement e.g. teacher is taking a reading/literacy/maths group and the rest of the class are working independently but are not directly supervised.
A6	Group work – movement	Learners working in groups, sitting at tables with some talking and movement e.g. working on projects or group discussion.

Table 17: Learning activities, description of activity and codes used in the noise surveys.

During each visit, a survey sheet was completed that included a drawing of the room and a record of the number of occupants. The Datalogging timecode from the sound level meter was also made along with the activity code, a brief description of the activity and the occurrence of any unexpected noise events. (see Appendix 14 for a copy of the noise survey template).

5.2.2 Data Analysis

After each visit, the data from the Casella 620A Integrated Digital Sound Level Meter was downloaded and saved as a .csv file. Each file was allocated a number which was recorded on the noise survey sheet. Data on L_{Aeq} and L_{Amin} was extracted and imported into Microsoft Excel 2016 where columns were created for each of the learning activities. The number of learners in the class was also added. Using the description of the learning activity and codes from the noise survey sheets, the L_{Aeq} and L_{Amin} data was categorised. The duration of each activity was also recorded. A single noise level figure for each subject and room was provided by arithmetically averaging the collated learning activities for each lesson. The overall noise level for each learning activity was obtained by arithmetically averaging the different activity codes.

A noise level-duration graph was created for each of the 150 lessons to illustrate the variation in noise over time. The graph shows how often each learning activity was observed during the lesson in two-minute timeframes. As Figure 29 illustrates, during the numeracy lesson in the control classroom three learning activities were observed: A3, A5, and A6. The least recorded activity was A5, observed for eight minutes at different points during the lesson. In contrast, group work with movement (A6) was the most observed activity, happening for a total of thirty-minutes at different time points during the lesson (refer to Appendix 15 for the noise level graphs from the control and intervention classrooms).

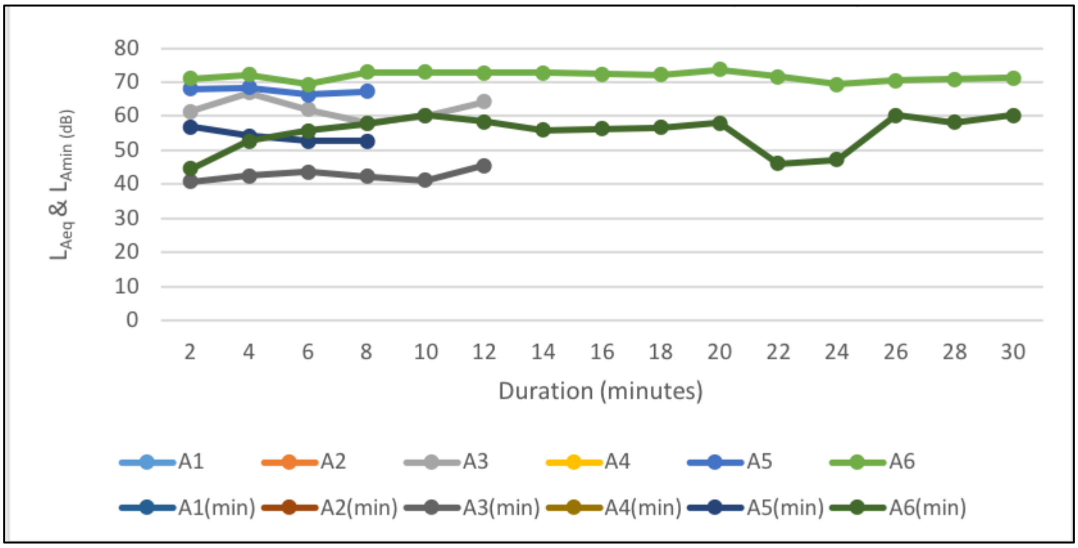


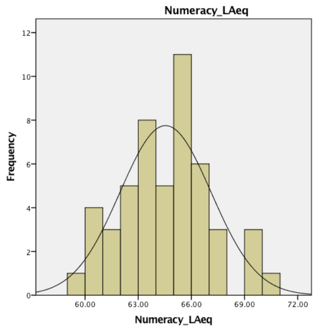
Figure 29: Variation of noise over time by learning activity. Each two-minute time point represents how often the activity was observed at different times in the lesson. This example is from a numeracy lesson in a control classroom.

The completed data for each classroom was imported into IBM SPSS version 23 for Mac where additional variables were created for control/intervention, school location, room volume, ambient noise levels, T_{mf}, C_{50}, SIMD quintile and the number of learners in the classroom. Descriptive statistics were used to summarise the duration and types of learning activities observed during lessons in numeracy, literacy, and IDL. For each lesson and learning activities, the L_{Aeq} and L_{Amin} means were calculated. Inferential statistics were used to determine if there was a significant difference in noise levels between the control and intervention classrooms.

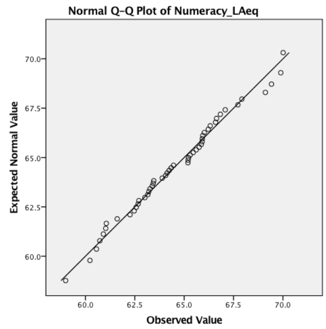
5.2.3 Distribution of the data

Field (2013) recommends testing the assumption of normality on data distribution through a process of visual inspection of the histogram, Q-Q-plots (Quantile-Quantile plots) and the conversion of any abnormality in skewness and kurtosis values into *z*-scores. The Q-Q-plots compares two quantiles against each other. As Figure 30 illustrates, the observed quantiles of the data (small circles) are compared against the quantiles expected if the data were normally distributed (straight diagonal line). Normally distributed data should fall along the straight diagonal line (Field, 2013). Figure 30 shows the histograms and Q-Q-plots from the noise surveys across all classrooms in the subject areas of literacy, numeracy, and IDL. Visual inspection indicates that there is negative skewness in both literacy and IDL which indicates that the majority of the noise levels were at the medium to higher end with only a smaller number at the lower end. To determine if the distribution of the data was normal the values of skewness and kurtosis were calculated.

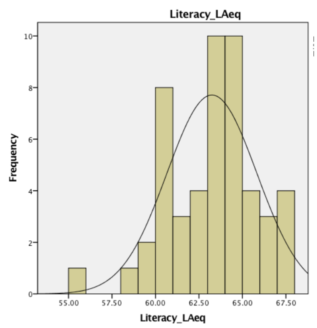
Numeracy histogram



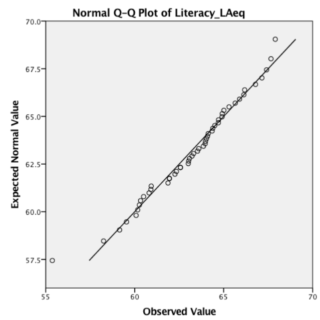
Numeracy Q-Q Plot



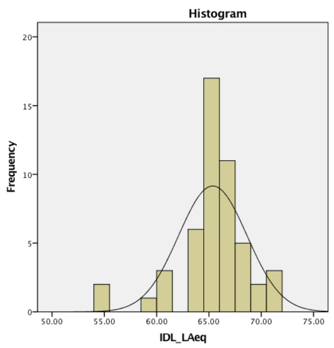
Literacy histogram



Literacy Q-Q Plot



IDL histogram



IDL Q-Q Plot

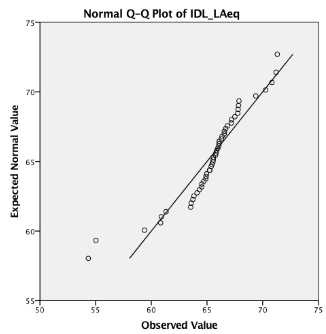


Figure 30: Histograms and Q-Q Plots from the occupied classroom noise surveys in numeracy, literacy and IDL.

In a normal distribution of data, values of skewness and kurtosis should be zero and the further the value is from this point the more likely the data is not parametric (Field, 2013). Hae-Young (2013) recommends obtaining z-scores by dividing skew and kurtosis values by their standard error using the following formula:

$$Z = \frac{Skew\ value}{SE_{skewness}}, \qquad Z = \frac{Excess\ Kurtosis}{SE_{excess\ kurtosis}}$$

To reject the null hypothesis and conclude the distribution of the data is non-normal depends on the sample size. For small samples ($n<50$) then z-scores for either skewness or kurtosis larger than 1.96 (with significance set at $p<0.05$) would reject the null hypothesis and for medium samples ($50<n<300$) the value would be 3.29 (with significance set at $p<0.05$) (Hae-Young, 2013). Fifty noise surveys were completed in each subject and so the null hypothesis was based on 1.96 (with significance set at $p<0.05$). As one of the aims of the noise surveys was to determine if there was a significant difference in noise levels across the three subject areas in the control and intervention classrooms z-scores were calculated for each condition by subject and all classrooms by subject, the results are presented in Table 18.

Classrooms	Skewness	$SE_{skewness}$	z-score	Kurtosis	$SE_{kurtosis}$	z score
All classes						
Literacy	-0.529	0.337	-1.57	0.553	0.662	0.84
Numeracy	0.114	0.337	0.34	-0.302	0.662	0.46
IDL	-1.340	0.337	-3.98	3.602	0.662	5.44
Control						
Literacy	-0.559	0.472	-1.18	0.327	0.918	0.36
Numeracy	-0.226	0.472	-0.48	-0.223	0.918	-0.24
IDL	-2.359	0.472	-4.74	8.109	0.918	8.83
Intervention						
Literacy	-0.200	0.456	-0.04	-0.972	0.887	-0.45
Numeracy	0.312	0.456	0.68	-0.399	0.887	-1.09
IDL	-1.080	0.456	-2.37	2.411	0.887	2.72

Table 18: Skewness, Kurtosis, and z-scores from the occupied classroom noise surveys.

The results show that the overall data is negatively skewed for literacy. The control group is more skewed than the intervention group. However, the z-scores indicate that this is not statistically significant and so there is no reason to reject the null hypothesis for either the combined data or condition. The numeracy data is evenly distributed and again there is no reason to reject the null hypothesis. The ILD z-scores indicate significant skewness and kurtosis values. Positive values in kurtosis are associated with a heavy-tailed distribution and this is apparent in the histograms in Figure 30. For the IDL data, the null hypothesis is rejected as the data distribution is non-normal.

In addition to the visual and numerical methods of testing the normality assumption of data distribution, it is also prudent to use a formal normality test. Yap and Sim (2011) compared four formal tests of normality and found that the Shapiro-Wilk test to be the most powerful. Table 19 provides the statistics and p-values for the three different subjects by all classes and conditions.

Classrooms	Statistics	p-value
All classes		
Literacy	0.976	0.393
Numeracy	0.984	0.707
IDL	0.880	0.001
Control		
Literacy	0.969	0.631
Numeracy	0.980	0.889
IDL	0.774	0.001
Intervention		
Literacy	0.983	0.930
Numeracy	0.946	0.183
IDL	0.919	0.043

Table 19: **Results from the Shapiro-Wilk test of normality show that the data for literacy and numeracy are normally distributed. IDL is not normally distributed.**

The results confirm the data from the histograms, Q-Q-plots, and z-scores that the data is normally distributed, apart from IDL. Therefore, a *t*-test (α=0.05) was used to compare the noise levels by curricular area and classroom activity in literacy and numeracy and a Mann-Whitney U test was used for IDL.

5.2.4 Duration of noise surveys by subject

The overall time for each noise survey by subject and activity is presented in Table 20. A total of 6014 noise survey minutes were recorded during the 150 lessons. In total, 2832 minutes (49.61%) were recorded in the control classrooms with 2876 minutes (50.39%) in the intervention classrooms. The amount of time that each curricular area was recorded in both the control and intervention classrooms was evenly spread.

	Control	Intervention	Total
Total survey in minutes (%)	2928 (48.7)	3086 (51.3)	6014
Literacy in minutes (%)	972 (49.4)	996 (50.6)	1968
Numeracy in minutes (%)	940 (49.0)	980 (51.0)	1920
IDL in minutes (%)	1016 (47.8)	1110 (52.2)	2126

Table 20: Descriptive statistics on the total time each curricular subject was surveyed in the control and intervention classrooms

Descriptive statistics on the length of the noise surveys in each subject are presented in Table 21. As seen, there is a slight variation in the mean and median values across all conditions and subject areas. This was mainly because unlike in secondary schools where there are rigidly set times for each lesson period, primary classes do not have rigorously enforced time allocations.

	Control			Intervention		
	Literacy	**Numeracy**	**IDL**	**Literacy**	**Numeracy**	**IDL**
Mean	40.5	39.2	42.3	38.3	37.7	42.7
St. Error	2.02	2.38	2.21	2.2	1.5	1.99
Median	40	37	45	36	36	42
Range	32	42	38	44	26	34
Minimum	26	26	26	26	26	26
Maximum	58	68	64	70	52	60

Table 21: Descriptive statistics showing the mean, median and range of noise surveys completed for all the control and intervention classrooms.

5.2.5 Duration of noise surveys by learning activities

Table 22 presents the percentage of time each learning activity was observed during lessons. Although the CfE recommends more active and interactive learning the results indicate that group work was the least used activity in literacy and numeracy across both the control and intervention classrooms. The most common observed teaching methodology in literacy was A5 and this was primarily as a result of the way reading groups are managed within schools. Class teachers and support staff will often take individual reading groups whilst the rest of the class complete other tasks individually, this was observed in both the control and intervention classrooms.

	Control						Intervention					
Activity Code	**A1**	**A2**	**A3**	**A4**	**A5**	**A6**	**A1**	**A2**	**A3**	**A4**	**A5**	**A6**
Literacy (%)	0	8	14	7	16	4	0	4	11	15	15	6
Numeracy (%)	0	5	14	12	13	6	0	5	11	14	13	7
IDL (%)	1	4	10	12	14	7	0	4	10	13	13	12

Table 22: Percentage of time each learning activity was observed during lessons in literacy, numeracy, and IDL.

In numeracy, activities associated with more than once person talking were the most commonly observed. The whole class would often be involved in reciting numeracy games, sometimes copying an audio recording. Group work was again not commonly observed. The majority of activities in both literacy and numeracy lessons involved more than one person talking, a degree of fidgeting and movement and these were evenly spread across the control and intervention classrooms. There is no published data on teaching styles in numeracy and literacy lessons to establish external validity.

IDL is a key component of the Scottish curriculum and is intended to promote learning within and across the eight different curricular areas. There is a broad range of learning experiences associated with IDL including collaborative learning, problem-solving and action-based research. The pedagogical methodology is less formal, exploratory with the focus on learner-centred rather than teacher-led learning (Humes, 2013). Therefore, it is not surprising that two-thirds of the activities observed involved movement and activity. Learning activities A2-A5 are fairly evenly observed across the control and intervention groups. There was a higher percentage of group work observed in the intervention classrooms.

5.2.6 Variation in noise levels by subjects

The mean L_{Aeq} and L_{Amin} noise levels are presented in Table 23. To determine whether the noise levels in literacy lessons were significantly different between the control and intervention groups two-tailed independent *t*-tests (α=0.05) were conducted on the L_{Aeq} and L_{Amin} values. The L_{Aeq} results show there was not a significant difference between the control (M=62.79, SD=2.95) and intervention (M=63.67, SD=2.18) conditions; $t(50)=-1.21$, p=0.23. The L_{Amin} data was tested using the same method and the results also indicated that there was not a significant difference between the control (M=42.43, SD=4.92) and intervention (M=43.76, SD=2.95) classrooms; $t(50)=-1.18$, p=0.246. The mean difference for L_{Aeq} was -0.88 and for L_{Amin} was -1.34.

School	Condition	N	Mean (L_{Aeq})	SD	Mean (L_{Amin})	SD
School 1	**soundfield**	2	63.69	0.93	39.73	0.70
	control	2	62.48	3.14	42.50	1.99
School 2	**soundfield**	2	63.76	1.05	45.44	2.00
	control	2	63.79	0.27	44.34	2.46
School 3	**soundfield**	2	62.74	1.12	44.98	0.28
	control	2	63.70	0.55	44.96	0.93
School 4	**soundfield**	2	62.57	1.59	45.56	1.08
	control	2	64.55	1.57	43.84	0.11
	control	2	62.15	0.28	43.92	1.53
School 5	**soundfield**	2	66.31	2.25	47.47	0.93
	control	2	65.07	1.57	45.10	2.57
School 6	**soundfield**	2	62.05	2.79	41.54	2.25
School 7	**soundfield**	2	64.91	0.01	45.26	0.36
	control	2	62.33	2.16	42.98	1.41
School 8	**soundfield**	2	64.55	1.89	41.27	2.39
School 9	**control**	2	57.26	2.67	31.17	1.98
	soundfield	2	63.32	1.51	42.00	5.02
School 10	**soundfield**	2	66.21	1.70	47.56	0.79
School 11	**soundfield**	2	63.16	5.12	41.92	3.87
	control	2	60.33	0.23	42.62	0.60
School 12	**soundfield**	2	61.10	1.09	43.60	0.28
	control	2	65.90	2.50	50.18	1.99
School 13	**soundfield**	2	63.29	3.30	42.63	3.67
	control	2	63.52	3.60	39.75	2.43
	control	2	62.70	6.29	37.79	6.71

Table 23: Mean and standard deviation noise levels (LA_{eq} and LA_{min}) observed during literacy lessons.

Further two-tailed independent *t*-tests (α=0.05) were carried out to determine whether there was a significant difference in L_{Aeq} and L_{Amin} levels in numeracy between the control and intervention classrooms. The mean L_{Aeq} and L_{Amin} noise levels for all twenty-five classrooms are presented in Table 24. In keeping with the findings from the literacy surveys, there was not a significant difference in L_{Aeq} scores between the control (M=64.78, SD=1.60) and the intervention (M=64.32, SD=3.24) classrooms; $t(50)=0.64$, $p=0.525$. The minimum noise levels were also not significantly different between the control (M=44.38, SD=3.38) and the intervention (M=44.04, SD=3.19) classrooms; $t(50)=0.37$, $p=0.714$. The overall mean L_{Aeq} and L_{Amin} noise levels in numeracy were higher than for the lessons in literacy. This was partly a consequence of whole-class activities that involved repeating numeracy rules or tables sometimes in conjunction with audio recordings.

School	Condition	N	Mean (L_{Aeq})	SD	Mean (L_{Amin})	SD
School 1	**soundfield**	2	63.48	1.80	42.00	3.16
	control	2	64.42	1.48	43.64	4.91
School 2	**soundfield**	2	62.84	0.24	42.07	0.74
	control	2	66.46	1.79	47.15	1.63
School 3	**soundfield**	2	63.94	2.37	45.86	0.41
	control	2	64.56	0.94	46.26	1.71
School 4	**soundfield**	2	64.32	2.29	47.34	0.83
	control	2	65.76	0.65	44.81	1.48
	control	2	64.83	3.19	47.74	4.16
School 5	**soundfield**	2	69.71	0.42	49.81	0.85
	control	2	64.65	1.73	43.83	2.92
School 6	**soundfield**	2	61.17	0.62	42.21	0.95
School 7	**soundfield**	2	62.89	3.30	43.33	2.06
	control	2	63.78	0.80	43.61	1.37
School 8	**soundfield**	2	63.47	3.61	43.29	3.20
School 9	**control**	2	63.83	3.91	42.65	4.86
	soundfield	2	65.54	3.40	44.85	0.56
School 10	**soundfield**	2	67.90	2.80	47.19	5.09
School 11	**soundfield**	2	67.97	1.61	39.98	1.90
	control	2	63.86	0.84	45.08	2.84
School 12	**soundfield**	2	59.61	0.87	43.31	0.44
	control	2	65.23	1.54	43.09	6.70
School 13	**soundfield**	2	63.26	1.12	41.33	1.67
	control	2	65.44	1.70	43.38	3.26
	control	2	64.61	1.63	39.77	3.72

Table 24: Mean and standard deviation noise levels (LA_{eq} and LA_{min}) observed during numeracy lessons.

The only two assumptions for completing the Mann-Whitney test is that the groups must be independent and that the dependent variable is either ordinal or numerical (continuous) (Brink, 2010). Both of these conditions were met for the IDL data. The mean L_{Aeq} and L_{Amin} noise levels for IDL in all the classrooms are presented in Table 25. The Mann-Whitney U test shows that there was not a significant difference

(U=345.5, p=0.515) in the L_{Aeq} noise levels between the control and intervention classrooms. The results for the L_{Amin} also show that there was not a significant difference between the two groups (U=263.5, p=0.346). The noise levels in IDL were higher than in numeracy and literacy. As many of the observed learning opportunities in this subject involve movement and discussion this was not surprising.

School	Condition	N	Mean (L_{Aeq})	SD	Mean (L_{Amin})	SD
School 1	**soundfield**	2	65.77	2.90	42.76	0.94
	control	2	65.81	2.39	46.01	4.83
School 2	**soundfield**	2	66.10	2.51	44.92	4.15
	control	2	63.71	0.86	43.87	0.78
School 3	**soundfield**	2	65.79	1.26	48.19	2.30
	control	2	65.12	0.58	45.59	1.30
School 4	**soundfield**	2	64.95	0.59	43.76	1.90
	control	2	63.30	3.49	44.89	3.10
	control	2	66.74	0.64	48.41	0.74
School 5	**soundfield**	2	71.25	0.78	51.07	3.46
	control	2	68.11	1.80	49.73	1.50
School 6	**soundfield**	2	62.97	5.06	42.18	4.13
School 7	**soundfield**	2	59.89	7.85	40.86	2.55
	control	2	60.13	7.21	43.10	6.67
School 8	**soundfield**	2	67.76	3.56	41.82	5.24
School 9	**control**	2	66.22	0.18	42.22	0.88
	soundfield	2	66.38	2.00	41.19	0.26
School 10	**soundfield**	2	68.75	1.40	46.16	3.73
School 11	**soundfield**	2	66.05	1.60	45.85	3.58
	control	2	65.94	0.19	46.31	2.56
School 12	**soundfield**	2	62.32	2.00	44.59	6.67
	control	2	64.55	1.44	44.82	2.00
School 13	**soundfield**	2	64.58	4.60	44.13	5.60
	control	2	65.29	1.08	45.45	3.25
	control	2	66.20	0.49	43.61	4.71

Table 25: Mean and standard deviation noise levels (LA_{eq} and LA_{min}) observed during IDL lessons.

The influence on room volume and T_{mf} on the L_{Aeq} levels in numeracy, literacy, and IDL were explored. Investigating all twenty-five classrooms using Spearman's correlation illustrated there was a significant correlation between the room volume and T_{mf} levels ($r=0.55$; $p<0.001$). The research classrooms with the highest ceilings also had the longest T_{mf}. This is consistent with previous studies that show a correlation between increases in RT and room volume (Shield et al., 2015).

5.3 Variation in noise levels by learning activities

The mean L_{Aeq} and L_{Amin} levels for each of the six learning activities are presented in Table 26. A heat diagram format has been used to illustrate the variation in noise levels by activity. As would be expected, the level of noise is higher for group work or activities involving movement and talking (‘chatter and clatter’) than for lessons involving more sedentary activities. In numeracy there is a 10.4 and 11 dB (A) L_{Aeq} difference between the quietest (A2) and noisiest (A6) activity in both the control and intervention conditions. In literacy, the difference was 9.7 dB (A) L_{Aeq} in the control and 12.1 dB (A) L_{Aeq} in the intervention classrooms.

	L_{Aeq}		L_{Amin}	
	Control	**Intervention**	**Control**	**Intervention**
Numeracy A1				
Numeracy A2	59.5	58.9	38.2	38.7
Numeracy A3	62.1	62.3	40.3	41.7
Numeracy A4	65.6	64.2	44.4	44.1
Numeracy A5	67.9	66.4	49.5	48.9
Numeracy A6	69.9	69.9	51.6	48.9
Literacy A1				
Literacy A2	57.6	57.2	38	38.6
Literacy A3	61.6	61.9	41.3	40.7
Literacy A4	65.8	64.8	45.8	44.7
Literacy A5	65.4	64.7	47.5	45.9
Literacy A6	67.3	69.3	47.2	51.8
IDL A1	48.5		33.7	
IDL A2	58.5	57.6	39.3	37.7
IDL A3	61.1	62	41.5	39.6
IDL A4	66.4	67	47.2	46.9
IDL A5	68.3	69.1	48.7	48.7
IDL A6	68.5	70.7	50.2	51.3

Table 26: Heat diagram illustrating the variation in noise levels by learning activity with the red palette representing higher L_{Aeq} and L_{Amin} levels and yellows/greens the lower level.

The activity commonly associated with the least intense noise levels (A1) was observed only once in IDL and this was a whole class test. The difference between the quietest and noisiest activity in IDL was 20 dB (A) $L_{Aeq.}$ This replicates the results from research in London primary schools where the difference between the quietest and noisiest learning activities was also 20 dB (A) L_{Aeq} (Shield and Dockrell, 2004). The highest average noise level was 70.7 dB (A) L_{Aeq} in IDL which is lower than 77 dB (a) L_{Aeq} observed in previous surveys in urban primaries (Shield and Dockrell, 2004). However, as there is a correlation between class numbers and noise levels this may be a consequence of this study design having a higher proportion of classrooms with a capped learner to teacher ratio of 1:18 (Shield and Dockrell, 2004, Dockrell and Shield, 2006). The L_{Amin} noise levels increase in parallel with the L_{Aeq} levels as the activity codes get higher (A1 to A6). This suggests that the minimal noise levels in the classrooms are influenced by the classroom occupants and the type of teaching methodology being used. This may be of interest to education policymakers as the CfE advocates a less formal teaching style with collaborative and active learning (Henderson, 2010).

In all three curricular areas, there was a negligible difference between the noise levels in learning activities A2-A3. These are activities that are generally associated with minimal amounts of movement.

5.3.1 Summary – occupied classroom noise surveys

As part of the process of establishing internal validity, 150 noise surveys were undertaken in core curricular areas in all the control and intervention classrooms. The overall mean noise levels were higher in IDL across both control and intervention classrooms. This was primarily due to the types of learning associated with this subject. Numeracy lessons had a higher mean L_{Aeq} value than literacy lessons. Data analysis was performed by subject using both parametric and non-parametric tests and the results show that there was not a significant difference in the L_{Aeq} and L_{Amin} noise levels by subject The survey of learning activities also demonstrate that activities associated with talking and movement have a higher noise level than those involving didactic teaching methodologies. Slack and Draugalis (2001) note that the process of establishing internal validity involves examining the study design and execution to discount alternative causes of the observed effects. The results from the noise surveys indicate that noise levels are not a situational variable that is significantly different across the two conditions.

5.4 Learners' listening in noise questionnaires

In addition to noise surveys, the views of the participants were sought to determine if there was a different experience of noise between learners exposed to dynamic soundfield amplification and the control group. The method used to capture the learners' experience of noise in the classrooms was a listening in noise questionnaire that was developed to be accessible to young people.

5.4.1 Data Collection method

The questionnaire was administered as a whole class exercise in each control and intervention classroom. The questionnaire was completed in the summer term to allow the intervention classrooms to become accustomed to the dynamic soundfield technology. Where learners were absent, return visits were made and these were completed on an individual or small group basis. This ensured that each learner completed all the questions and so eliminated attrition. Bryman (2015) observes that missing data in questionnaires arise when the participant either accidentally overlooks a question or does not want to answer it. To avoid missing data the students received instructions to complete each question one at a time and these were then checked before attempting the next question. As well as the researcher, staff from the Sensory Support (Deaf Learners) Service and class teacher invigilated this process. Where a student had not completed a question, an additional explanation was provided. As additional insurance against unanswered questions, the completed questionnaires were then checked after the visit, and a return appointment made where necessary. This method ensured that all the questions were completed by the participants.

5.4.2 Data analysis

All the questionnaire data was inserted into IBM SPSS version 23 for Mac where each participant is regarded as an individual case and assigned a unique ID. The questionnaire was constructed using a closed set of questions with each participant either required to make a binary choice from a list of predefined options or rate different situations using a Likert scale. Variables were created for each of the possible answers and allocated a code. Table 27 presents the different variables measured, codes and recodes, tools used and the primary method of analysis.

Variable	Variable Type	Tool	Variable Codes	Analysis
Sources of noise	Categorical	Learner questionnaire (Q.1)	1 = Yes 2 = no	Chi-square Cramer's V
Noise levels	Categorical	Learner questionnaire (Q.2)	1 = low 2 = medium 3 = high	Chi-square Cramer's V
Emotions	Categorical	Learner questionnaire (Q.3)	1 = Yes 2 = no	Chi-square Cramer's V
Hear the teacher	Categorical	Learner questionnaire (Q.4)	**Original variable** 1 = Always Easy 2 = Mostly Easy 3 = Not Easy or hard 4 = Mostly hard 5 = Always hard **Recoded variable** 1 = easy 2 = not easy	Chi-square Logistical regression

Table 27: Variables, codes and the statistical methods used to analyse the questionnaires.

The questionnaire was divided into four parts: awareness of classroom noise, levels of classroom noise, feelings towards noise and ability to hear the teacher. The purpose of question 1 was to determine if there was a significant difference between the learners' awareness of environmental and occupant generated noise in the control and intervention classrooms. Three different noise categories are generally observable in classrooms: noise external to the school, noise inside the school but external to the classroom and internal classroom noise (Department of Education and Skills, 1993). The polar responses were coded 1 when the learner circled an answer and 0 for a non-response. Question 2 was designed to measure the learners' perception of noise levels during ten different curricular areas to determine if there was a significant difference between the two conditions. As Table 27 illustrates noise levels were coded low (1), medium (2) and high (3). Question 3 was intended to gain an insight into the feelings the individual learners had towards the effects of noise on their well-being. Seven complementary antonyms were provided, and the learners were instructed that they could choose as many as were appropriate. The answers were sub-categorised into three groups: emotions, attention, and fatigue. The learners could select from five positive and five negative emotions. Responses for attention and fatigue were a binary negative and positive choice. The overall purpose of question 3 was to find if there was a significant difference in the attitudes of learners towards noise in the control and intervention classrooms.

Question 4 was designed to see if there was a significant relationship between the ability to hear the teacher easily in noise and exposure to dynamic soundfield technology. Six different scenarios were presented, and the participants responded using a five-point Likert scale: always easy, mostly easy, not hard or easy, mostly hard and always hard. This was based on a modified version of the Listening Inventory for Education (LIFE) which was developed for use with deaf learners (Anderson, 1998). The LIFE was not appropriate for this study as it is long, with fifteen different questions in total. Furthermore, there is no reference to the noise levels associated with each learning activity and as such the full listening context is not given. Each of the six-classroom scenarios was allocated a noise category based on the noise levels observed during six learning activities identified by Shield and Dockrell (2008), a modified version is presented in Table 17. Learning activities associated with sitting and one person talking (56.3 – 61.2 L_{Aeq}) were categorised as low. Noise levels associated with movement and group work (72.2-76.8 L_{Aeq}) were categorised as high (Shield and Dockrell, 2008). Four of the classroom activities were associated with high noise levels were two activities were categorised as low.

As discussed in Chapter 2, visual cues such as speech-reading and gestures can mitigate the effects of masking by noise (Cherry, 1953). The questionnaire was developed to measure the experience of learners when the face of the teacher was visible and so speechreading was accessible and when speechreading was inaccessible. Four further scenarios are presented: group work (A6), students walking around the room (A5), Students talking (A5) and teacher walking (A4). As Table 27 illustrates, the 5-point Likert scale used to record the participants'' responses were recoded using SPSS into binary categories (see section 5.7).

5.4.3 Methods of statistical analysis

The primary form of analysis was the Chi-square test with a significance level of α=0.05. The test was selected as it not only provides information on the significance of any observed differences but also details the frequency of any observations for each category compared to the expected frequency that would be observed by chance (McHugh, 2013). If the observed values are greater than the expected values (α=0.05), the null hypothesis can be rejected. To measure the strength of the difference between the expected and observed values in each cell the adjusted residuals were explored. The adjusted residuals are the difference between the expected and observed values in each cell divided by the standard deviation of all the residuals using the following formula (Bewick et al., 2004):

$$d_{ij} = \frac{O_{ij} - E_{ij}}{\sqrt{E_{ij}\left(1 - \frac{n_{i.}}{N}\right)\left(1 - \frac{n_{.j}}{N}\right)}}$$

O_{ij} and E_{ij} are the residuals (observed and expected), n_i is the total frequency of the row, n_j is the total frequency of the column and N is the total number of observations. The adjusted residuals are comparable to a *z*-score and so if the residuals are greater than ± 2 then they are significant. The larger the difference from ± 2 the more significant the association between the two variables (Bewick et al., 2004).

To determine whether to accept or reject the null hypothesis a Chi-square analysis was performed on the questionnaire results from the control and intervention classrooms. Since one of the aims of the research was to measure whether the attainment gap was reduced between learners from the most and least deprived areas, analysis of the responses from learners in SIMD 1 and 5 quintiles both between and within the control and intervention groups was performed. Figure 31 illustrates the methods used to analyse the questionnaire data.

In addition to the Chi-square test, question 4 was analysed using binary logical regression. Regression analysis is a method of predicting an outcome variable from one or more predictor variables (Field, 2013). This is discussed in section 5.9.

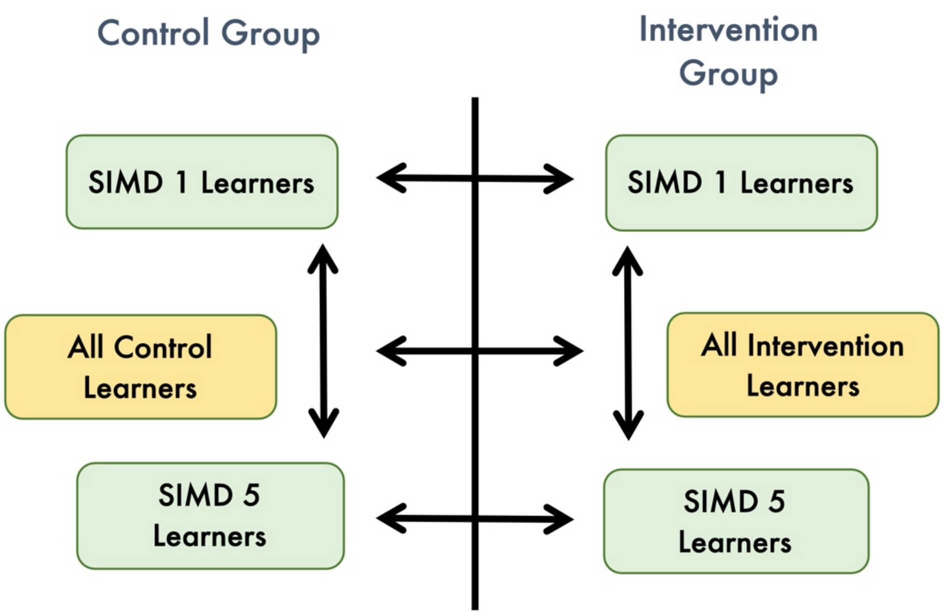


Figure 31: Chi-square analysis of the questionnaire results.

5.4.4 Effect Size

In Chi-square, when a variable has more than two levels or the contingency table is large the Cramer V is a common method used to test the strength of the association if a significant result is observed. Cramer V is calculated using the following formula where n=the number of rows or columns, whichever is less (McHugh, 2013):

$$\sqrt{\frac{X^2/n}{(k-1)}} = \sqrt{\frac{x^2}{n(k-1)}}$$

However, caution needs to be applied to the results as one of the criticisms of Cramer V is its tendency to produce low correlation measures when there are highly significant results (McHugh, 2009).

5.4.5 External Noise

The null and alternative hypothesis for external noise is:

H_0: There is no difference in the awareness of external noise sources between learners in a control classroom and an intervention classroom fitted with a dynamic soundfield system.

H_1: There is a difference in the awareness of external noise sources between learners in a control classroom and an intervention classroom fitted with a dynamic soundfield system.

To establish the sources of noise pollution in the control and intervention classroom the learners were able to select from twenty-four different internal and external noise sources commonly heard in school classrooms. Three external noise sources were presented in the questionnaire along with one choice that could be from either external or internal sources. Table 28 illustrates the most commonly occurring external noise originated from the playground. The prevalence of playground noise may be a consequence of some schools having split lunchtimes between the lower and upper school and consequently the playground area is occupied during these times. Furthermore, parents generally drop-off and pick up young people from school and congregate in the playground area. Previous noise studies outside 142 London primary schools found that road traffic was the most common external noise with it recorded outside of 86% of schools (Shield and Dockrell, 2004). However, the current study design excluded schools that were located on the main road with all the schools either sitting on a side street or set back from the main street and so sheltered from the effects of road traffic. In addition, many of the research schools had large grounds which also helped to cushion the effects of traffic noise.

External noise	Control	Intervention	Total	χ^2 Values
Planes and trains	61	67	128	(1, N=495)=1.02, p=0.31
Road Traffic	82	84	166	(1, N=495)=3.14, p=0.78
Noise form the playground	179	213	389	(1, N=495)=2.55, p=0.11
Other?	127	168	295	(1, N=495)=1.84, p=0.67

Table 28: Results from the noise questionnaire showing the number of learners reporting external noise in the control and intervention classrooms along with the combined total.

As Table 28 illustrates there was not a significant difference between the control and intervention rooms on the four external noise categories. There is no reason to reject the null hypothesis. Dockrell and Shield (2012) also found that the installation of soundfield equipment had no significant impact on learners' awareness of external sources of sound. Further analysis was run to explore if there was a difference between learners from the most and least deprived quintiles in the control and intervention classroom. There was no significant difference for SIMD 1 learners: planes and trains [χ^2 (1, N=155)=0.05, p=0.83], road traffic [χ^2 (1, N=155)=0.13, p=0.72], playground [χ^2 (1, N=155)=0.27, p=0.60] and other [χ^2 (1, N=155)=1.56, p=0.21]. A similar outcome was observed for SIMD 5 learners: planes and trains [χ^2 (1, N=183)=1.4, p=0.24], road traffic [χ^2 (1, N=183)=1.7, p=0.19], playground [χ^2 (1, N=183)=1.4, p=0.23] and other [χ^2 (1, N=183)=0.41, p=0.52]. There is no reason to reject the null hypothesis.

An analysis of the responses for learners from quintiles 1 and 5 within the control and intervention classrooms was also performed. The analysis indicated that within the intervention classrooms there was a significant difference between learners from SIMD 1 compared to SIMD 5 on the awareness of road traffic noise [χ^2 (1, N=192) =4.75, p=0.029]. Fifty-seven point nine per cent of learners from SIMD 1 were aware of road traffic noise compared to 42.1% for SIMD 5. An examination of the adjusted residuals shows that for both SIMD 1 (z=2.2) and SIMD 5 (z=-2.2) the strength of the difference between the expected and observed values were significant. There was also a significant difference for other unspecified noises for SIMD 5 learners compared to SIMD 1 [χ^2 (1, N=192) =3.99, p=0.046]. 60.2% of SIMD 5 learners identified this as a noise source compared to 39.8% of SIMD 1 learners. An examination of the adjusted residuals reveals that for both SIMD 1 (z=-2) and SIMD 5 (z=2) the strength of the difference between the observed and expected frequency in each cell was significant. Within the control group, learners from SIMD 5 were significantly more aware of playground noise compared to SIMD 1 [χ^2 (1, N=146) =4.72, p=0.03]. 58.5% of learners from SIMD 5 were aware of playground noise compared to 41.5% for SIMD 1. The strength of the difference between the expected and observed values were significant: SIMD 1 (z=-2.2) and SIMD 5 (z=2.2). The Cramer V had a small effect size for all external noise sources.

5.4.6 Internal to the school, external to the classroom

The null and alternative hypothesis for internal noise to the school but external to the classroom is:

H_0: *There is no difference between the awareness of internal noise to the school but external to the classroom and learners in the control and intervention classrooms fitted with a dynamic soundfield system.*

H_1: *There is a difference between the awareness of internal noise to the school but external to the classroom and learners in the control and intervention classrooms fitted with a dynamic soundfield system.*

The questionnaire identified six internal noise sources that were external to the classroom along with one undefined choice. As Table 29 illustrates the main source of noise originated from the occupants of the school. The noise pollution either came from other classrooms or was created by the movement of learners through the school. Seventy-nine point four per cent of learners across both conditions identified noise from other classrooms and seventy-eight point eight per cent identified movement around the school.

Table 29 illustrates the noise from the corridor and other classrooms were significantly different between the control and intervention classrooms. An examination of the adjusted residuals shows the strength of the difference was significant for corridor noise: control (z=3.3) and intervention classrooms (z=-3.3). The Cramer V strength of the association was weak (0.150). The adjusted residuals for noise from other classrooms were also significant in both conditions: control (z=2.2) and intervention (z=-2.2). Once again, the Cramer V was weak (0.1). The results suggest that a higher proportion of learners in the control classrooms were aware of both noise sources compared to the intervention.

Internal to the school but external to the classroom	Control	Intervention	Total	χ^2 Values
Noise from toilet	71	72	143	(1, N=495)= 2.76, p=0.97
Noise from people in the corridors	186	204	390	(1, N=495)= 11.09, p<0.001
Noise from other classrooms	182	211	393	(1, N=495)= 4.73, p=0.03
Other noises – unknown	117	144	261	(1, N=495)= 0.22, p=0.64
Noise from dinner hall	10	12	22	(1, N=495)=0.24, p=0.88
Music	131	184	315	(1, N=495)=1.78, p=0.18
Noise from the gym hall	16	18	34	(1, N=495)=0.15, p=0.70

Table 29: Results from the noise questionnaire showing the number of learners reporting noise external to the classroom but internal to the school, in the control and intervention classrooms along with the combined total.

Further analysis was carried out to explore the difference between learners from SIMD quintiles 1 and 5 and the seven noise categories in the control and intervention classrooms. SIMD 1 learners' awareness of noise from other classrooms [χ^2 (1, N=155) =9.09, p=0.03] and noise from the corridor [χ^2 (1, N=155) =10.69, p<0.001] was significant. The strength of the difference for noise from other classrooms was significant: control (z=3) and intervention (z=-3). The Cramer V of 0.242 indicates there was a small association. The adjusted residuals for corridor noise was also significant: control (z=3.3) and intervention (z=-3.3). The strength of the association was again small (0.263). These results indicate that SIMD 1 learners in the control condition were more aware of noises from the corridors and other classrooms than their counterparts in the intervention classrooms.

An examination within the control classrooms found that learners from SIMD 1 were more aware of noise from other classrooms than SIMD 5 learners, [χ^2 (1, N=146) =5.32, p=0.021]. The strength of the difference was significant: SIMD 1 (z=2.3) and SIMD 5 (z=-2.3). The Cramer V was small (0.191). In contrast, there was no significant difference between SIMD 1 and 5 learners in the intervention classroom on the awareness of noise from other classrooms [χ^2 (1, N=192) =0.525, p=0.469]. Noise from people in the corridors was significantly different between learners from SIMD 1 and 5 within the control classroom [χ^2 (1, N=146) =6.41, p=0.011]. Once again, there was no significant difference in the intervention classrooms [χ^2 (1, N=192) =0.145, p=0.704]. The strength of the difference between the observed and expected frequency in each cell was significant for SIMD 1 (z=2.5) and 5 (z=-2.5). The Cramer V indicated that the strength of the association was low (0.210). The results suggest that SIMD 1 learners in the control classrooms were more affected by the noise from the corridor and other classrooms than their peers in the classrooms exposed to dynamic soundfield.

5.4.7 Internal classroom noise

The null and alternative hypothesis for internal classroom noise is:

H_0: There is no difference between the awareness of internal classroom noise and learners in the control and intervention classrooms fitted with a dynamic soundfield system.

H_1: There is a difference between the awareness of internal classroom noise and learners in the control and intervention classrooms fitted with a dynamic soundfield system.

Thirteen different internal noise sources were presented in the questionnaire which can be categorised into three groups: classroom equipment, building services, and occupant generated. As Table 30 illustrates, in general, classroom equipment was heard more often than noise from building services with interactive whiteboards and computers being most prevalent. Noise from lights was the most commonly observed building service. However, the most prominent noise source across both the control and intervention classrooms was from the learners themselves. Active learning, students' movement, and talking were all identified as the main sources of noise. As Table 30 shows, there are no significant differences between the control and treatment groups on the different sources of internal noise. There is no reason to reject H_0.

Internal to the classroom	Control	Intervention	Total	χ^2 Values
Noise from fan	31	37	68	(1, N=495)=0.98, p=0.75
Noise from Computer	113	154	267	(1, N=495)=0.54, p=0.46
Buzzing from light	61	62	123	(1, N=495)=2.20, p=0.14
Art lessons	166	227	393	(1, N=495)=1.98, p=0.16
Golden Time	183	229	420	(1, N=495)=0.34, p=0.56
Children talking near me	193	255	448	(1, N=495)=1.10, p=0.29
Children talking in other parts of the class	174	237	411	(1, N=495)= 2.2, p=0.14
Noise from printer	49	56	105	(1, N=495)= 0.43, p=0.51
Smartboard Projector	104	131	235	(1, N=495)=0.32,p=0.86
Children getting things from their bag	101	117	220	(1, N=495)=0.98, p=0.32
Scraping of chairs and table leg	171	208	381	(1, N=495)=1.07, p=0.29
Clattering pens and pencils	171	227	398	(1, N=495)= 0.63, p=0.43
Heater	31	32	63	(1, N=495)= 0.85, p=0.36

Table 30: Results from the noise questionnaire showing the number of learners reporting sources of internal noise in the control and intervention classrooms along with the combined total.

Further between-subject analysis was performed on SIMD 1 and 5 participants and it similarly found no difference between the control and intervention groups. Once again, there is no reason to reject the null hypothesis.

Comparisons were made between participants from SIMD 1 and 5 within the control and intervention classrooms. All internal noise sources were analysed apart from the noise from heaters as the Chi-square assumption that each cell should have a minimum of five was not met. In the control classroom, noise generated by movement and fidgeting significantly affected learners from the most socially deprived areas more than those from SIMD 5: clattering of pens and pencils [χ^2 (1, N=146) =11.43, p<0.001], learners getting things from their bags [χ^2 (1, N=146) =19.26, p<0.001] and scraping of chairs and table legs [χ^2 (1, N=146) =11.96, p<0.001]. An exploration of the cells found that learners from SIMD 1 were significantly above the expected values and SIMD 5 learners below. The strength of the difference between the observed and expected frequency in each cell was significant: getting things from their bags SIMD 1 (z=4.4) and 5 (z=-4.4); scraping of chairs and table legs SIMD 1 (z=3.5) and 5 (z=-3.5) and pens and pencils SIMD 1 (z=3.4) and 5 (z=-3.4). In all cases the Cramer V strength of the association was small.

In the dynamic soundfield classrooms, SIMD 1 participants were also primarily affected by the noise created from fidgeting and movement with scraping of chairs and table legs [χ^2 (1, N=192) =15.12, p<0.001] and learners getting things from their bags [χ^2 (1, N=192) =9.50, p=0.02] identified as significant. The strength of the difference between the observed and expected frequency in each cell was significant: getting things from their bags SIMD 1 (z=3.1) and 5 (z=-3.1); scraping of chairs and table legs SIMD 1 (z=3.9) and 5 (z=-3.9). The strength of the association was once again small.

Looking at the results two common trends are apparent. The first is that young people from SIMD 1 were more affected by intermittent noise created by people moving either around the school or from other classrooms compared to learners from SIMD 5. Furthermore, learners from SIMD 1 in the control classroom were more affected by noise internal to the school but external to the classroom than those exposed to dynamic soundfield. The second trend is that intermittent noise created by people getting things from bags and scrapping chairs affected SIMD 1 learners more than SIMD 5 in both the control and intervention classrooms.

5.5 Noise levels by subjects

The null and alternative hypothesis for classroom noise is:

H_0: There is no difference in the awareness of classroom noise levels during lessons in the control and intervention classrooms fitted with a dynamic soundfield system.

H_1: There is a difference in the awareness of classroom noise levels during lessons in the control and intervention classrooms fitted with a dynamic soundfield system.

The individual learners in the control and intervention classrooms were asked to assess the noise levels in ten curricular areas using a three-point Likert scale. The students were asked to circle a scaled thermometer that was colour coded green, orange and red which represented three categories of noise levels: low, medium and high. The ten subjects were selected as they represent common curricular activities in primary schools. The traffic light colour coding was selected as primary schools use this system as part of the peer and self-assessment process and so was familiar to the participants (Hallam et al., 2004). The primary aim was to establish if there was a significant difference between perceived noise levels during lessons and exposure to dynamic soundfield technology.

The results are presented in Table 31. Unsurprisingly, in practical subjects or activities involving movement, a higher number of respondents reported noise levels to be high. Noise levels are generally higher when there are group activities that involve moving around the classroom or multiple people talking such as Art. In subjects such as spelling and circle time that are commonly associated with sitting and listening, there was a higher proportion of low-level noise responses. Core subjects such as writing and mathematics were considered to have medium noise levels by the majority of learners in the control and intervention classrooms. The perceptions of noise by the participants are generally consistent with the results from the occupied classroom noise surveys. A Chi-Square test of independence on all ten subject areas indicated that there was no significant difference between the noise levels in the classes exposed to the soundfield intervention and the control group. There are no grounds to reject the null hypothesis.

	Control				Intervention				
Subject	Low	Medium	High	Total	Low	Medium	High	Total	χ^2 Values
Writing	43	127	47	217	65	157	56	278	(2,N=495)=0.93, p=0.63
Maths	45	93	79	217	55	145	78	278	(2,N=495)=4.93, p=0.80
Circle time	124	63	30	217	181	72	25	278	2,N=495)=4.26, p=0.12
Spelling	93	83	41	217	94	125	59	278	2,N=495)=4.27, p=0.12
Project work	37	96	84	217	44	127	107	278	2,N=495)=0.17, p=0.92
Problem solving	79	65	73	217	88	111	79	278	2,N=495)=5.31, p=0.70
Reading	87	70	60	217	110	95	73	278	2,N=495)=0.23, p=0.89
Golden Time	8	30	179	217	10	29	239	278	2,N=495)=1.36, p=0.51
Show and Tell	121	81	15	217	136	117	25	278	2,N=495)=2.29, p=0.32
Art	36	68	113	217	33	88	157	278	2,N=495)=2.38, p=0.30

Table 31: Perceived levels of noise in ten curricular areas in the control and intervention classrooms.

Between-subject analysis on the participants from SIMD 1 and 5 quintiles was performed and this demonstrated there was a significant difference in the perceived levels of noise during lessons in mathematics and spelling. In mathematics, this was only observed by SIMD 5 learners' [χ^2 (2, N=183) =16.28, p<0.001]. In the control classrooms, 63.1% of SIMD 5 learners reported noise levels to be high during lessons in mathematics compared to 36.9% in the intervention. The difference was significant: control (z=4) intervention (z=-4). The Cramer V strength of the association was low, 0.298. In spelling, SIMD 5 learners also recorded significant differences [χ^2 (2, N=183) =6.63, p=0.036]. Seventy-five per cent of learners in the intervention group observed noise levels to be high for spelling compared to 25% in the control. The adjusted residuals were significant: control (z=-2.3) and intervention (z=2.3). The Cramer V value of 0.190 shows that the strength of the association was low.

Within the control and intervention groups analysis was performed based on social deprivation. In the control classroom, mathematics was again identified by SIMD 5 learners as having significantly higher noise levels (68.3%) in comparison to SIMD 1 learners (31.7%): [χ^2 (2, N=146) =8.79, p=0.012]. The strength of the difference between the observed and expected values in each cell was significant: SIMD 1 (z=-2.9) and 5 (z=2.9). The Cramer V was small. In the intervention classrooms, the noise levels during problem-solving were also perceived differently by SIMD 1 and 5 learners [χ^2 (2, N=183) =6.28, p=0.043]. A higher percentage of SIMD 5 learners observed noise levels to be high (63.2%) compared to SIMD 1 (36.8%). The strength of the difference in each cell was significant: SIMD 1 (z=-2) and 5 (z=2). The strength of the association was low (0.181).

Lessons in Show and Tell generally involve young learners sitting and listening to their peers present on a topic of interest. There was a significant difference in the perception of noise by SIMD 1 and 5 learners in the intervention classrooms. Once again SIMD 5 learners reported noise levels to be high (73.8%) compared to SIMD 1 (26.32%): [χ^2 (2, N=192) =20.12, p<0.001]. However, the number of learners that made this observation was small: SIMD 1 (n=5), SIMD 5 (n=14). The strength of the difference was significant: SIMD 1 (z=-2.8) and 5 (z=2.8). A higher percentage of SIMD 1 observed that noise levels were low (58.7%) during Show and Tell compared to SIMD 5 (41.3%). The adjusted residuals show that the difference was significant SIMD 1 (z=3.4) and 5 (z=-3.4). More participants from SIMD 5 observed noise levels to be medium (64.2%) compared to SIMD 1 (35.8%) and again the adjusted residuals were significant: SIMD 1 (z=-2.4) and 5 (z=2.4). The Cramer V was small to moderate (0.324).

Looking at the results two common trends are apparent. Firstly, for the majority of subjects, there was not a significant difference between the learners' perception of noise in the control and intervention classrooms. Secondly, SIMD 5 learners perceived noise levels in certain subjects to be higher than SIMD 1 learners. Between subject analysis found that learners from the 20% least deprived areas perceived noise levels to be higher in mathematics and spelling compared to learners from the most deprived areas. Within the control group, SIMD 5 learners observed noise levels to be higher in mathematics than SIMD 1 learners. In the intervention classrooms, this trend was also observed during problem-solving and show and tell activities.

5.6 Feelings towards noise

The null and alternative hypothesis for feelings towards noise is:

H_0: There is no difference in the learners' feelings towards noise in the control and intervention classrooms fitted with a dynamic soundfield system.

H_1: There is a difference in the learners' feelings towards noise in the control and intervention classrooms fitted with a dynamic soundfield system.

The results of the participants' feelings towards the noise are presented in Table 32. Unsurprisingly, across both the control and intervention classrooms the learners generally identified emotions that showed dissatisfaction with noise. Disliking the noise and frustration were the most commonly observed emotions. Several of the learners also responded with indifference towards the classroom noise. A Chi-square test analysis was performed to determine if there were different emotional responses towards noise in the control and intervention classrooms. As Table 32 illustrates, the results were non-significant for all the emotions indicating there were no differences between the two conditions. Further analysis of the participants from SIMD quintiles 1 and 5 both between and within the group also found that there was no significant difference. There is no reason to reject the null hypothesis.

	Emotions	Soundfield	Control	Total	χ^2 Values
happy	Observed	72	62	134	(1,N=495)=0.44,p=0.51
	Expected	75.3	58.7	134	
like the noise	Observed	47	37	84	(1,N=495)=0.002,p=0.97
	Expected	47.2	36.8	84	
calm	Observed	69	55	124	(1,N=495)=0.18,p=0.89
	Expected	69.6	54.4	124	
not bothered	Observed	115	83	198	(1,N=495)=0.49,p=0.48
	Expected	111.2	86.8	198	
excited	Observed	51	40	91	(1,N=495)=0.001,p=0.98
	Expected	51.1	39.9	91	
dislike the noise	Observed	152	125	277	(1,N=495)=0.42,p=0.52
	Expected	155.6	121.4	277	
frustration	Observed	159	114	273	(1,N=495)=1.07,p=0.31
	Expected	153.3	119.7	273	
sad	Observed	125	102	227	(1,N=495)=0.20,p=0.65
	Expected	127.5	99.5	227	
mad	Observed	144	103	247	(1,N=495)=2.11,p=0.35
	Expected	138.7	108.3	247	
grumpy	Observed	116	91	207	(1,N=495)=0.002,p=0.96
	Expected	116.3	90.7	207	

Table 32: Emotions expressed by the participants in the control and intervention classrooms towards the noise.

The participants in both the control and intervention classrooms were asked about how well they could focus in noise. Most of the learners (n=325) in both the control and intervention classrooms found it difficult to focus in noise but there was not a significant difference between the two conditions [χ^2 (1, N=495) =2.03, p=0.15]. Although 61.5% of participants exposed to dynamic soundfield responded that they were able to focus in noise compared to 38.5% in the control this again was non-significant [χ^2 (1, N=495) =2.07, p=0.15]. Between-group analysis of SIMD 1 [χ^2 (1, N=155) =0.93, p=0.34] and 5 [χ^2 (1, N=183) =0.39, p=0.84] learners also found there was not a significant difference between the control and intervention classrooms. Analysis within the control [χ^2 (1, N=146) =0.33, p=0.57] and intervention [χ^2 (1, N=192) =0.24, p=0.13] groups also found no significant difference between learners from SIMD 1 and 5.

The participants were also asked to record if the noise made them tired. Fifty-four point two per cent in the intervention classroom reported feeling tired compared to 45.8 per cent in the control. Once again there was not a significant difference [χ^2 (1, N=495) =0.871, p=0.35]. Between-group analysis also found no significant difference for SIMD 1 [χ^2 (1, N=155) =0.008, p=0.93] and SIMD 5 [χ^2 (1, N=183) =0.40, p=0.53] learners. There was also no significant difference between SIMD 1 and 5 learners in the control [χ^2 (1, N=146) =0.36, p=0.55] and intervention classrooms [χ^2 (1, N=192) =0.18, p=0.89].

5.7 Hearing the teacher

The null and alternative hypothesis for hearing the teacher is:

H_0: There is no difference in accessing the spoken words of the class teacher between learners in the control and intervention classrooms fitted with a dynamic soundfield system.

H_1: There is a difference in accessing the spoken words of the class teacher between learners in the control and intervention classrooms fitted with a dynamic soundfield system.

The learners' responses to the six questions about hearing the teacher are presented in Table 33. For question 1, the key assumptions of a Chi-Square Test that each cell should have a minimum value of five is not met. Field (2013) recommends that if any cell does not meet the minimum value that either more data is collected or try to increase the proportion of cases falling into each category. The original data was collected and coded using a five-point Likert scale that ranged from always easy (1) to always hard (5). One of the strengths of the Likert scale for perception or opinion-based questionnaires is that it provides more opportunity for the participant to provide sensitive and subtle information compared to a dichotomous format. Cohen (2018) rightly observes that the interpretation of a five-point scale can be problematic as it can be difficult for the participants to distinguish the nuance between values of always/mostly hard and always/mostly easy. For young people, it cannot always be assumed that a linear scale with equal values between each of the five semantic rating points (Always Easy, Mostly Easy, Not Easy or Hard, Mostly Hard, Always Hard) is applied. To address this issue and to increase the number of cases in each cell the five-point Likert scale was summated and recoded into two variables in SPSS version 23 for a Mac: a positive response, hear the teacher easily (coded 1) and a negative response, not hear the teacher easily (coded 2). The responses always easy (1) and mostly easy (2) were recoded 1 (hearing the teacher easily*)*. The responses mostly hard (4) and always hard (5) were recoded 2 (not hearing the teacher easily).

Dealing with the mid-point (not hard or easy) in a Likert scale is also challenging as it could suggest that the respondent is ambivalent or neutral on the topic. Cohen (2018) advises that the golden rule is applied: crude data produces a crude interpretation and subtle data provides subtle interpretation. As discussed in

Chapter 2, listening to spoken language in a classroom is a complex process that not only involves identification of the spoken word, storage, and processing of information but also applying knowledge and previous experience to the educational task in hand. Learners have finite cognitive capacity and when the primary task (listening) becomes more demanding due to an imperfect listening environment performance on the secondary task may depreciate (Choi et al., 2008, Gosselin and Gagné, 2010). The further you move from an optimal listening condition (Always Easy) the more listening effort is required. In an optimal listening environment, the consumption of additional cognitive resources should be minimal and so there will be little listening effort. In contrast, complex listening environments will be effortful and challenging. The mid-point response 'not hard or easy' (3) infers the consumption of some cognitive resources on the part of the listener and as such was recoded 2.

How well can you hear the teacher?		Always Easy	Mostly Easy	Not Easy or Hard	Mostly Hard	Always Hard	Total
			Intervention				
Q1 You can see the face; noise levels are low	Observed	206.0	41.0	22.0	7.0	2.0	278
	Expected	189.3	41.0	33.1	11.2	3.4	278
Q2 You cannot see the face; noise levels are low	Observed	135.0	85.0	38.0	11.0	9.0	278
	Expected	117.9	73.6	51.1	21.9	13.5	278
Q3 The teacher is walking around the room, noise levels are high	Observed	93.0	80.0	74.0	23.0	8.0	278
	Expected	84.8	68.5	73.6	34.8	16.3	278
Q4 Other students are talking; noise levels are high	Observed	92.0	62.0	52.0	50.0	22.0	278
	Expected	77.5	55.6	52.8	50.5	41.6	278
Q5 Other students are walking; noise levels are high	Observed	95.0	78.0	61.0	29.0	15.0	278
	Expected	85.9	68.5	59.0	36.5	28.1	278
Q6 Working in a group; noise levels are high	Observed	100.0	66.0	50.0	40.0	22.0	278
	Expected	84.8	59.0	56.7	43.2	34.3	278

How well can you hear the teacher?		Always Easy	Mostly Easy	Not Easy or Hard	Mostly Hard	Always Hard	Total
			Control				
Q1 You can see the face; noise levels are low	Observed	131.0	32.0	37.0	13.0	4.0	217
	Expected	147.7	32.0	25.9	8.8	2.6	217
Q2 You cannot see the face; noise levels are low	Observed	75.0	46.0	53.0	28.0	15.0	217
	Expected	92.1	57.4	39.9	17.1	10.5	217
Q3 The teacher is walking around the room, noise levels are high	Observed	58.0	42.0	57.0	39.0	21.0	217
	Expected	66.2	53.5	57.4	27.2	12.7	217
Q4 Other students are talking; noise levels are high	Observed	46.0	37.0	45.0	49.0	40.0	217
	Expected	60.5	43.4	42.5	43.4	27.2	217
Q5 Other students are walking; noise levels are high	Observed	58.0	44.0	44.0	36.0	35.0	217
	Expected	67.1	53.5	46.0	28.5	21.9	217
Q6 Working in a group; noise levels are high	Observed	51.0	39.0	51.0	37.0	39.0	217
	Expected	74.1	42.1	44.7	31.1	25.0	217

Table 33: Responses from the learners in the control and intervention classroom on how well they could hear the teacher in six different teaching scenarios.

The recoding of variables presented in Table 27 could introduce bias as the researcher is not blind to the outcome. To determine if there was a significant difference between the experience of hearing the teacher easily in the control and intervention classrooms a Chi-Square test was performed on both the original and recoded data. Thereafter, to establish if bias was introduced an examination of the p-values from the original and recoded was undertaken. The results in Table 34 show that for each question there was a significant difference between the control and intervention classrooms on hearing the teacher easily in the six different classroom scenarios. There was no change in the significance values ($p<0.001$) between the recoded and original data for the five questions that were compared. Although the Chi-Square test shows an association between dynamic soundfield technology and hearing the teacher it does not indicate whether other variables in the data set also contribute to this outcome. Therefore, logistical regression was calculated.

How well can you hear the teacher?		Easily	Not Easily	Recoded χ^2 Values	Original χ^2 Values
You can see the face	Intervention	247	31	(1,N=495)=16.16, p=<0.001	Not tested
	Control	163	54		
You cannot see the face	Intervention	220	58	(1,N=495)=31.07, p=<0.001	(4,N=495)=33.12, p=<0.001
	Control	121	96		
The teacher is walking around	Intervention	173	97	(1,N=495)=15.11, p=<0.001	4,N=495)=24.97, p=<0.001
	Control	100	120		
Other students are talking	Intervention	154	124	(1,N=495)=14.36, p=<0.001	(4,N=495)=20.17, p=<0.001
	Control	83	134		
Other students are walking	Intervention	173	105	(1,N=495)=11.44, p=<0.001	(4,N=495)=22.76, p=<0.001
	Control	102	115		
Working in a group	Intervention	166	112	1,N=495)=16.23, p=<0.001	(4,N=495)=20.50, p=<0.001
	Control	90	127		

Table 34: Recoded and original χ^2 values on how well the learners in the control and intervention classrooms could hear the teacher.

5.8 Logistical Regression

Logistic regression was selected as the method of analysis as it is an established technique for research in which there is a binary dependent variable and two or more independent (predictor) variables. The predictor variables can be a combination of continuous and categorical variables (Reed and Wu, 2013). There are several assumptions applicable to the logistical regression model. The first is that the errors need to be of independent, so any observations should not be matched or repeat measures. A between-subject design is also presumed. The assumption was met as the questionnaire was completed on a single occasion, near the end of the study by the participants in both the control and intervention classrooms.

A large sample size is also required, as there is a relationship between the sample size and the calculation of the odds ratio. The size of the odds ratio is inversely related to the sample size and so a reduced sample size increases the bias and overestimation in the odds ratio produced. To minimise this effect small sample sizes should be avoided with a minimum of 200 and a figure of around 500 being proposed as an appropriate sample size (Nemes et al., 2009, Reed and Wu, 2013). Another method to calculate the appropriate sample size for regression analysis is based on the number of predictor variables. Green (1991) applied the formula $N > 104 + m$ when testing individual predictors with a medium effect size, where m is the number of predictor variables. VanVoorhis and Morgan (2007) advise that when the model contains six or more predictors, a minimum of ten participants per predictor is appropriate. There were thirteen predictor variables in the regression model and so the 495 participants in the current study meet the sample size assumption.

5.8.1 Multicollinearity

Another assumption of logistic regression is that there should not be a perfect relationship between any of the predictor variables. Multicollinearity occurs when predictor variables within the study are correlated to each other. Predictor variables that show multicollinearity will adversely affect the results in a logistical regression model as they inflate variances and so incorrect inferences could be made about the relationship between variables. There is a consensus that diagnostic assessment of multicollinearity is best performed using linear regression and measuring the condition index, variance inflation factor (VIF), tolerance and variance proportions. In addition, the Pearson correlation coefficient can also be used to screen for multicollinearity (Midi et al., 2010, Field, 2013).

As a rule of thumb, any correlation coefficient greater than 0.8 indicates that there is an issue with multicollinearity (Midi et al., 2010). The correlation coefficient between the variables T_{mf} and C_{50} (0.80) signifies that there are concerns of multicollinearity. Further analysis was performed using the collinearity statistics tolerance and VIF. There is not a consensus in the literature regarding the cut off value to use when identifying multicollinearity with VIF: Field (2013) uses the value of ten and Midi et al. (2010) argue that for models using logistical regression values above 2.5 should be a concern. As a binary logistic model was part of the study design any values above 2.5 was regarded as a concern. As Table 35 illustrates both T_{mf} and C_{50} indicate multicollinearity. In addition to VIF, tolerance levels also provide information on multicollinearity and although there is no formal cut off value anything below 0.1 is generally regarded as a cause for concern. There were no concerns identified.

Model	**Tolerance**	**VIF**
Gender	0.978	1.022
ASN	0.903	1.107
Age (decimal)	0.944	1.059
Ethnic groups	0.870	1.150
Languages of the home	0.766	1.306
Language level	0.776	1.289
Attendance	0.848	1.179
Room volume	0.633	1.581
T_{MF}	0.275	3.631
Ambient noise	0.860	1.163
C_{50}	0.323	3.100
SIMD	0.674	1.484

Table 35: Detection of multicollinearity in the predictor variables.

A diagnostic assessment for multicollinearity also requires the examination of the eigenvalues and the condition index. The condition index is the square root of the ratio of the largest eigenvalue to the eigenvalue of interest. Generally, as eigen values decrease the condition index will increase, eigenvalues close to zero indicate significant issues with multicollinearity. A condition index of around fifteen is generally a cause for concern, anything greater is a serious concern. Examination of the variance proportions also gives supplementary details on the amount of variance each predictor contributes to the eigenvalue. Any predictors that have high proportions on the same small eigenvalue would indicate collinearity (Midi et al., 2010). The largest condition index was 69.87 with the eigenvalue of 0.002. The variance proportions show that 73% of the variance in the regression coefficient for room volume and T_{mf} is related to this.

Midi et al. (2010) suggests that if there is not the option of increasing the sample size then omitting all but one of the correlated predictor variables can reduce multicollinearity. Therefore, room volume and T_{mf} were removed from the logistic regression model.

5.8.2 Linearity of the Logit

Another assumption of logistic regression is linearity. Although there is no assumed linear relationship between the independent and dependent variables, the logit of the outcome and predictor variable values require to be related linearly (Field, 2013). There are four continuous predictor variables in the questionnaire regression model: C_{50}, ambient noise levels, age, and school attendance. Four dummy (log) variables were created in SPSS and these were forced into a single block binary regression model. As part of this process. the four dummy log variables interact with their original variable counterpart. Any interaction that is significant (<0.05) will mean that the assumption of linearity of the logit has been violated. All values were >0.05 and so the assumption of linearity has been met. Furthermore, an examination of the standardised residuals identified no outliners.

5.8.3 Predictor variable selection

The predictor variables were selected based on the known barriers to listening in noise discussed in Chapter 2. The categorical predictor variables included: gender, ethnicity, English as an additional language, language competency, ASN, SIMD and dynamic soundfield (seven categorical variables). In addition, the following continuous predictors were included: age, speech clarity, ambient noise levels and school attendance (four continuous variables). Dichotomous dummy predictor variables were created for gender (1=male, 0=female), ethnicity (1=white, 0=other), English as an additional language (1=English, 0=other), language competency (1=fluent, 0=other), ASN (1=yes, 0=no) and dynamic soundfield (1=soundfield, 0=control). The multicategory variable SIMD has 5 (*k*) levels and these were recoded using the *k*-1 dummy variable method. The dependent variable was hearing the teacher easily (coded 1) and not hearing the teacher easily (coded 0).

One issue with having too many predictor variables in a regression model is that it can compromise the true associations, lead to large standard errors and imprecise confidence intervals (Ranganathan et al., 2017). The relevant predictors were selected using five automated fitting methods: entry, forward likelihood ratio, forward conditional, backward likelihood ratio and backward conditional. As the purpose of the process was to identify suitable predictor variables the cut-off for significance is generally higher ($p<0.1$) than the conventional $p<0.05$ (Aziz et al., 2016, Ranganathan et al., 2017). As Table 36 illustrates ten variables were identified as significant.

Variables	Regression entry method		
	Entry	**Forward LR/ Forward C**	**Backward LR/ Backward C**
Demographic Variables			
Age	-	-	Q1- 0.097**
Gender	- Q6 - 0.068**	- -	Q3 - 0.09** Q6 - 0.066**
Ethnicity	Q1 – 0.078**	Q1 - 0.035*	Q1 - 0.028*
English as an additional language	Q2 – 0.51** Q4 - 0.046* Q6 - 0.078**	Q2 - .017* -	Q2 - 0.017* -
Attendance	Q3 - 0.058**	-	Q3 - 0.081**
SIMD **1** **2** **5**	 Q2 - 0.04* Q2 - 0.067** Q2 - 0.025* -	 -	 Q5 - 0.069**
Environmental Variables			
Dynamic soundfield	Q1 - 0.001* Q2 - 0.001* Q3 - 0.001* Q4 - 0.001* Q5 - 0.001* Q6 - 0.001*	Q1 - 0.001* Q2 - 0.001* Q3 - 0.001* Q4 - 0.001* Q5 - 0.001* Q6 - 0.001*	Q1 - 0.001* Q2 - 0.001* Q3 - 0.001* Q4 - 0.001* Q5 - 0.001* Q6 - 0.001*
C_{50}	Q1 - 0.081** Q6 - 0.082**	- -	- Q6 - 0.078**

**= significance P<0.1 *= significance p<0.05 - not selected

Table 36: Predictor variables that were selected from the five regression fitting methods.

5.8.4 Model Fit

The regression model was built using the forced entry method and the overall fit and predictive accuracy was assessed using the likelihood ratio test. The 2-log likelihood assesses the overall fit of the two models, one with the constant and no predictor variables added ($-2LL_{null}$) and the other a saturated model with the *k* variables included ($-2LL_k$). The difference between the observed and fitted values provides a goodness of fit index ($2LL_{null}$ - $-2LL_k$) and a Chi-Square distribution, with *k* degrees of freedom. The difference in the value of the -2LL should be less than with only the constant in the model. The lower the value the more accurate the model (Aziz et al., 2016). The difference in the -2LL is tested against a Chi-Square distribution to give a *p*-value which indicates whether the model is better at significantly predicting the fit than the null model (Field, 2013).

Another method to measure the goodness of the fit is the Hosmer and Lemeshow R^2_L test. It again tests the null hypothesis that the model is fitted correctly and produces a *p*-value output. Using a significance level (α=0.05), a $p<0.05$ would suggest the model is not a good fit and conversely a $p>0.05$ would indicate the model is an acceptable fit. It is appropriate to test using the Hosmer and Lemeshow test as the Chi-Square test is effective when all the predictor variables are categorical and if the number of observations for each category is five. When there are continuous predictor variables or the number of cases per profile is below five then the *p*-values may be inaccurate. The overall fit of the model was assessed using the 2-log likelihood statistic, Chi-Square and the Hosmer and Lemeshow.

As Table 37 illustrates the saturated model is better at predicting the fit for all six questions than the null (constant) model. A good model of fit should have large Chi-Square values and significant *p*-values. The Hosmer and Lemeshow also shows that the model is a good fit as the p-values are all non-significant ($p>0.05$).

Questions	Log-Likelihood		Model Coefficient		Hosmer & Lemeshow
	$-2LL_{null}$	$-2LL_k$	Chi-square	Sig.	
Question 1	454.01	425.44	28.57	0.001	χ^2(8,N=495)=12.0, p=0.15
Question 2	613.79	566.48	47.304	0.001	χ^2(8,N=495)=2.99, p=0.94
Question 3	682.12	656.74	25.383	0.031	χ^2(8,N=495)=3.82, p=0.87
Question 4	685.33	652.53	32.797	0.003	χ^2(8,N=495)=13.47, p=0.097
Question 5	680.09	653.43	26.661	0.02	χ^2(8,N=495)=4.22, p=0.84
Question 6	685.63	643.33	42.30	0.001	χ^2(8,N=495)=4.17, p=0.84

Table 37: Overall fit of the regression model.

5.8.5 Effect Size - Odds Ratio and Risk Ratio

Long (2014) observes that meaningful interpretations from binary logistical regression is based on predicted probabilities, sometimes referred to as the risk ratio and the functions of those probabilities: odds and odds ratios. The odds ratio is a ratio of two odds; the odds that an event will occur divided by the odds that the event will not occur (Liberman, 2005). In experimental research involving control and intervention groups, the odds ratio is the odds that an event of interest will occur/not occur in the treatment group divided by the odds that an event of interest will occur/not occur in the control group. The following formula applies, where PG_1 represents the odds of an event occurring in the intervention group and PG_2 represents the odds of an event occurring in the control (McHugh, 2009).

$$\text{Odds ratio} = \frac{PG_1/(1 - PG_1)}{PG_2/(1 - PG_2)}$$

There is not a consensus on the effect size for the odds ratio. A number of authors have used the following standard: small=1.5, medium=2 large=3 (Sullivan and Feinn, 2012, Hae-Young, 2015). Chen et al. (2010) calculated the odds ratio equivalent to Cohen's $\boldsymbol{d}$ at a 5% significance level: $\boldsymbol{d}$ (0.2) small=1.52, $\boldsymbol{d}$ (0.5) moderate=2.74 and $\boldsymbol{d}$ (0.8) =4.74. Finally, Ferguson (2009) used a higher measure: small=2, moderate=3, and strong=4.

The odds ratio can be converted to probability using the following formula (Liberman, 2005):

$$p = \frac{\text{Odds}}{\text{Odds } + 1}$$

The risk ratio is also used to interpret the outcome of logistic regression. The risk ratio is calculated by dividing the risk in the intervention group by the risk in the control group (Sainani, 2011). In many studies, the risk ratio is incorrectly derived from the odds ratio calculation. When the outcome of interest is low then the risk ratio and odds ratio can be similar but when the outcome of interest is frequent the odds ratio will overestimate the risk ratio. Zhang and Yu (1998) devised the following formula to correct the odds ratio and provide a more accurate representation of risk:

$$\text{RR} = \frac{OR}{(1 - p_0) + (p_0 \times\ OR)}$$

P_0 is the outcome in the control group, RR is the estimated risk ratio and OR is the odds ratio. The effect size for the risk ratio is defined as follows: small=2, medium=3, and strong=4 (Ferguson, 2009, Sullivan and Feinn, 2012, Hae-Young, 2015).

5.8.6 Questions 1 & 2 – teacher at the front, noise levels low

Using the ability to hear the teacher easily as the dependent variable, the ten predictor variables were fitted into a binary logistical regression model using the entry selection method. In question one the teacher is standing at the front of the classroom; noise levels are low, and the listener has full access to visual cues to support the spatial release from masking. As Table 38 illustrates, the categorical variable dynamic soundfield (Wald=12.16, P<0.001) is the only significant predictor (α=0.05) for hearing the teacher easily. The odds ratio was 2.44 (95% CI 1.48-4.03) which suggests that in the intervention classroom learners were 2.44 times more likely to hear the teacher easily than in the control classrooms. Expressed as probability, it can be seen that learners in the intervention classroom are 71 per cent more likely to hear the class teacher easily than in the control. Using the various indices, this can be classified as a small to medium effect size. The risk ratio was calculated, where P_0=0.41 which gave a small risk ratio effect of 1.54. The learners in the control classroom were 1.54 times at risk of not hearing the teacher compared to the intervention group exposed to dynamic soundfield technology.

	S.E.	Wald	Sig.	Odds Ratio	95% C.I. for OR	
					Lower	Upper
Dynamic soundfield	0.26	12.16	0.001	2.44	1.478	4.027
Gender	0.25	0.25	0.62	1.13	0.543	1.439
Ethnicity	0.47	3.26	0.073	0.433	0.926	5.756
English as an Additional language	0.63	0.1	0.75	1.22	0.356	4.176
Attendance	0.02	0.02	0.89	1.01	0.957	1.051
SIMD1	0.40	0.03	0.87	1.06	0.427	2.065
SIMD2	0.43	1.42	0.23	1.66	0.261	1.387
SIMD5	0.39	0.13	0.72	0.869	0.403	1.874
C_{50}	0.07	2.73	0.099	1.126	0.978	1.296
Age	0.39	2.76	0.097	0.521	0.241	1.125

Table 38: Results from the logistic regression model for question 1 using ten predictor variables.

In question two the teacher is standing at the front of the room; noise levels are low but visual cues required to support the release from masking are not available to the listener. As Table 39 shows dynamic soundfield technology (Wald 29.90, $p<0.001$) is significantly associated with hearing the teacher easily in the intervention classrooms. The positive odds ratio of 3.15 (95% CI 2.08-4.75) indicates that in classrooms exposed to the dynamic soundfield learners were 3.15 times more likely to hear the teacher easily than in the control condition. This is a moderate effect size. Cherry (1953) identified visual cues as one of the discriminating features that allow the target signal of interest to be identified by the listener in an acoustically complex environment. It is interesting to note that the odds ratio is higher for question 2 than question 1, in which the only different variable is the ability to see the teacher's face. The probability that learners exposed to dynamic soundfield can hear the teacher more easily than the learners in the control condition is 75 per cent. The risk ratio was calculated, where P_0=0.32 giving a small risk ratio effect of 1.87.

Learners in the control condition are 1.87 times less likely to hear the teacher well compared to the intervention classrooms.

	S.E.	Wald	df	Sig.	Odds Ratio	95% C.I. for OR	
						Lower	Upper
Dynamic soundfield	0.21	29.90	1	0.001	3.15	2.09	4.75
Gender	0.20	0.05	1	0.818	1.05	0.64	1.42
Ethnicity	0.44	0.49	1	0.484	1.36	0.58	3.18
English as an additional language	0.56	2.79	1	0.095	2.56	0.85	7.68
Attendance	0.02	0.01	1	0.912	1.00	0.96	1.04
SIMD1	0.31	2.53	1	0.112	1.64	0.89	3.03
SIMD2	0.35	1.62	1	0.203	1.57	0.78	3.13
SIMD5	0.30	3.31	1	0.069	1.73	0.96	3.13
C_{50}	0.06	0.12	1	0.728	1.02	0.91	1.15
Age	0.32	0.55	1	0.458	0.79	0.42	1.48

Table 39: Results from the logistic regression model for question 2 using ten predictor variables.

5.8.7 Question 3 – teacher moving around, noise levels high

In question 3 the teacher is moving around the classroom and so limited visual cues are available of the listener. Unlike in question 2 the lack of visual cues is inferred rather than explicit. Noise levels are high which is consistent with learning activities A4-A6 from the classroom noise surveys. As Table 40 illustrates only one predictor variable was significant at the 5% level: learners exposed to dynamic soundfield technology (Wald=13.17, $p<0.001$). The odds ratio indicates that in the dynamic soundfield classrooms learners were 1.99 times (95% CI 1.37-2.90) more able to hear the teacher easily compared to the control. This is a small to medium effect size. Interestingly, this is a smaller effect size than when noise levels were lower in questions 1 and 2. Expressed as probability, the learners in the dynamic soundfield classroom were 67 per cent more likely to hear the teacher easily compared to the control classroom. Listeners in the control classroom are 1.34 times less likely to hear the teacher easily compared to the intervention classroom, this is a small risk ratio.

	S.E.	Wald	df	Sig.	Odds Ratio	95% C.I. for OR	
						Lower	Upper
Dynamic soundfield	0.19	13.17	1	0.001	1.99	1.37	2.90
Gender	0.19	2.79	1	0.095	1.34	0.51	1.06
Ethnicity	0.41	0.04	1	0.841	1.09	0.48	2.45
English as an additional language	0.54	0.27	1	0.604	1.32	0.46	3.83
Attendance	0.02	3.56	1	0.059	1.04	1.00	1.08
SIMD1	0.29	0.24	1	0.625	1.15	0.49	1.54
SIMD2	0.33	1.19	1	0.274	1.44	0.36	1.33
SIMD5	0.28	0.01	1	0.933	1.02	0.59	1.78
C_{50}	0.06	0.73	1	0.392	1.05	0.94	1.17
Age	0.30	0.03	1	0.867	1.05	0.58	1.90

Table 40: Results from the logistic regression model for question 3 using ten predictor variables.

5.8.8 Question 4 & 5 – learners talking and walking, noise levels high

In question 4 the teacher is speaking to the class and the other learners in the room are talking. This type of scenario was classified in the classroom noise survey as Activity A3. The mean noise levels recorded in the control classroom were 61.63 dB(A) L_{Aeq} and in the intervention classrooms 62.2 dB(A) L_{Aeq}. As Table 41 illustrates only dynamic soundfield (Wald=14.50, p<0.001) was positively associated with hearing the teacher easily. The odds ratio suggests that learners in the dynamic soundfield classrooms are 2.07 times (95% CI 1.42-3.01) more likely to hear the teacher easily compared to the control classroom. Expressed as probability, learners in the intervention classrooms are 67 per cent more likely to hear the teacher easily compared to the control group. The risk ratio (P_0=0.48) was small, with learners in the control condition 1.37 times less likely to hear the teacher easily compared to the intervention classrooms.

.

	S.E.	Wald	df	Sig.	Odds Ratio	95% C.I. for OR	
						Lower	Upper
Dynamic soundfield	0.19	14.50	1	0.001	2.07	1.42	3.01
Gender	0.19	1.67	1	0.196	1.27	0.55	1.13
Ethnicity	0.42	0.25	1	0.621	0.81	0.36	1.85
English as an additional language	0.58	3.15	1	0.076	2.81	0.90	8.78
Attendance	0.02	0.43	1	0.511	1.01	0.98	1.05
SIMD1	0.29	0.01	1	0.907	1.04	0.55	1.71
SIMD2	0.33	0.22	1	0.642	1.17	0.61	2.23
SIMD5	0.28	0.67	1	0.412	.970	0.73	2.18
C_{50}	0.05	0.00	1	0.970	1.00	0.90	1.12
Age	0.30	0.04	1	0.842	0.94	0.52	1.70

Table 41: Results from the logistic regression model for question 4 using ten predictor variables

Question 5 referred to learning activities A4 and A5 from the classroom noise survey, where other learners were talking and walking around the classroom. The mean noise levels in the control classrooms were (A4) 65.77dB(A) L_{Aeq} and (A5) 67.1dB(A) $L_{Aeq.}$ The intervention classrooms recorded mean noise levels of (A4) 65.57 dB(A) L_{Aeq} and (A5) 66.76 dB(A) L_{Aeq}. As Table 42 indicates dynamic soundfield (Wald=12.02, P<0.001), was the only significant (α=0.05) predictor variable. The odds ratio indicated that learners in the intervention classroom were 1.93 times (95% CI 1.33-2.81) more likely to hear the teacher easily compared to the control condition. This was a small effect size. Expressed as a probability, the learners in the intervention classroom were 66% more likely to hear the teacher easily. The risk ratio (P_0=0.52) was small, with learners in the control condition 1.30 times less likely to hear the teacher easily compared to the intervention classrooms.

	S.E.	Wald	df	Sig.	Odds Ratio	95% C.I. for OR	
						Lower	Upper
Dynamic soundfield	0.19	12.02	1	0.001	1.93	1.33	2.81
Gender	0.19	0.27	1	0.605	1.10	0.63	1.31
Ethnicity	0.41	0.09	1	0.760	1.13	0.51	2.53
English as an additional language	0.54	0.10	1	0.754	1.18	0.41	3.40
Attendance	0.02	0.04	1	0.838	1.00	0.96	1.03
SIMD1	0.29	0.98	1	0.322	1.33	0.43	1.32
SIMD2	0.33	0.42	1	0.518	1.24	0.42	1.54
SIMD5	0.28	0.00	1	0.951	1.01	0.57	1.71
C_{50}	0.06	1.32	1	0.250	0.94	0.84	1.05
Age	0.30	0.05	1	0.829	0.94	0.52	1.68

Table 42: Results from the logistic regression model for question 5 using ten predictor variables.

5.8.9 Question 6 – group work, noise levels high

Question 6 evaluated how easily learners could hear the teacher during group work. This was classified as learning activity A6 in the noise survey where the mean noise levels recorded in the control classrooms were 68.73 dB(A) L_{Aeq} and 69.88 dB(A) L_{Aeq} in the intervention classrooms. As Table 43 illustrates there is a positive association with hearing the teacher easily for the predictor variable dynamic soundfield (Wald=16.05, P<0.001). According to the odds ratio learners in the soundfield classrooms were 2.22 times more likely to hear the teacher easily (95% CI 1.53-3.24) compared to not hearing the teacher easily in the control classrooms. Expressed as a probability, the logistic regression model suggests that learners in the dynamic soundfield classrooms are 69 per cent more likely to hear the teacher easily compared to learners in the control environment. This is a small to medium effect size. The risk of learners (P_0=0.52) in the control classroom not hearing the teacher easily is 1.43, this is a small effect size.

	S.E.	Wald	df	Sig.	Odds Ratio	95% C.I. for OR	
						Lower	Upper
Dynamic soundfield	0.19	17.28	1	0.001	2.22	1.53	3.24
Gender	0.19	3.15	1	0.076	1.39	0.50	1.04
Ethnicity	0.42	0.09	1	0.762	0.88	0.38	2.01
English as an additional language	0.57	2.64	1	0.104	2.53	0.83	7.73
Attendance	0.02	0.51	1	0.473	1.01	0.98	1.05
SIMD1	0.29	0.67	1	0.413	1.27	0.45	1.39
SIMD2	0.33	1.14	1	0.285	1.43	0.37	1.34
SIMD5	0.28	0.00	1	0.945	1.02	0.59	1.77
C_{50}	0.06	2.05	1	0.152	0.92	0.83	1.03
Age	0.30	0.64	1	0.424	0.79	0.43	1.42

Table 43: Results from the logistic regression model for question 6 using ten predictor variables.

5.8.10 Summary – hearing the teacher

Exposure to dynamic soundfield technology was the only predictor variable that was significantly associated with hearing the teacher easily across all six different listening experiences. Regardless of the levels of background noise, the young learners in the dynamic soundfield classrooms reported that they could hear the teacher more easily than the control. Interestingly, question 2 had the highest odds ratio and this was when noise levels were low, but the face of the teacher was not visible. This would suggest that young people make use of visuals cues to mitigate the effects of masking in a complex listening environment. This aligns with research by Atilgan et al. (2018) which suggests that when there is temporal coherence, visual and auditory stimuli is bound together to create a multimodal object in the auditory cortex and this process supports auditory scene analysis. Overall, visual stimulation has the effect of boosting or amplifying auditory information. It would also appear that the dynamic soundfield provided an additional benefit to the listener when the visual cues that assist with the unmasking of speech are removed. The six different scenarios presented in the questionnaire covered many of the teaching methodologies in a classroom. The results suggest that dynamic soundfield was beneficial in all six areas.

5.9 Teacher questionnaire data collection

The 28 teachers who participated in this research completed a short questionnaire at the end of the intervention. There was no attrition. As teachers have a heavy workload, the survey was designed to provide the maximum amount of data with the minimum amount of demand. The questionnaire in the control classroom was a single page and the intervention classroom contained two pages (see Appendix 8). Teachers in both the control and intervention classrooms were asked to comment on the source and level of classroom noise. In addition, the teachers in the intervention classroom were asked to comment on the dynamic soundfield system.

As with any research measurement, response and expectation bias compromises internal validity. The threats in the current study were the teachers in the dynamic soundfield classrooms were not blind to the intervention. Furthermore, there was also an ongoing relationship between the teachers and the researcher throughout the study as the Datalogging from the dynamic soundfield system was collected at least every thirty days.

Although the results need to be treated with caution, at the end of a longitudinal study it is necessary to collect feedback on the intervention from the participants. As the answers could be subject to response and expectation bias, open-ended or semi-structured questions that would collect qualitative data on the experience of the teachers in the intervention classrooms were not appropriate. The questionnaire was composed of six closed questions, designed to measure the impact of the dynamic soundfield system using a 5-point Likert scale. The teachers had to delete or circle their responses.

5.9.1 Data analysis

The responses from the teacher questionnaires were inputted into IBM SPSS version 23 for Mac. As it was a closed questionnaire, with a limited number of participants the results were analysed using descriptive statistics. The results were compared to the data from the learner questionnaire to determine if there was agreement on the key areas of noise and the impact of the intervention.

5.9.2 Sources of noise

Teachers' awareness of different internal and external noise sources typically heard in the classroom was measured. As Figure 32 illustrates the only external noise sources recorded were from the playground; 21.4 per cent of the teaching staff (n=6) were aware of this form of noise pollution. This contrasts with 78.6 per cent (n=389) of primary 3 learners who recorded playground noise as the most prominent external noise. The results suggest that adults are better at ignoring external noise pollution compared to young learners in this research.

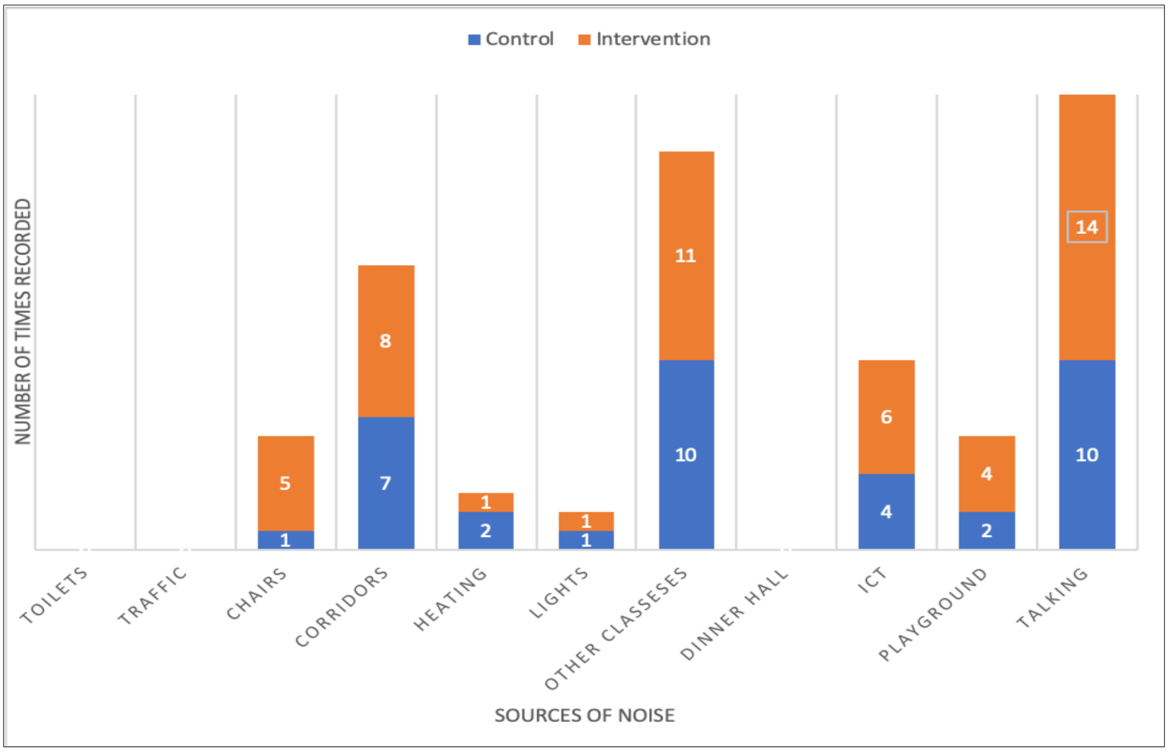


Figure 32: Sources of noise perceived by the class teachers in the control (blue) and intervention (orange) classrooms.

The most commonly observed noise from inside the school but external to the classroom was from other classrooms, 75 per cent of teachers (n=21) were aware of this noise in both the control and intervention classrooms. This percentage is similar to the 80.2 per cent (n=397) of primary 3 learners who also recorded this as the most commonly heard noise source. 53.6 per cent of teachers (n=15) were aware of noise from people in the corridor which was lower than the 78.8 per cent for primary 3 learners (n=390).

The internal classroom noise that was most prominent came from other learners talking. 85.7 per cent of the class teachers (n=24) recorded this. Only a small proportion of teachers, 21.4 per cent (n=6) were aware of the scraping of chairs, the majority of these were from the intervention classrooms. This contrasts sharply with the young learners where 77 per cent (n=381) were aware of this form of internal noise pollution.

5.9.3 Noise levels during lessons and activities

The class teachers were asked to rate the noise levels in their class from 1 (very low) to 5 (very high). Figure 33 shows that teachers in the control and intervention classrooms identified spelling as a subject that generally had lower noise levels and Golden Time had higher levels. Overall practical subjects were regarded as having higher noise levels, which is comparable to the results from the learner questionnaire. Overall, there was a fairly even spread of the noise levels by subject and activity across both the control and intervention classrooms.

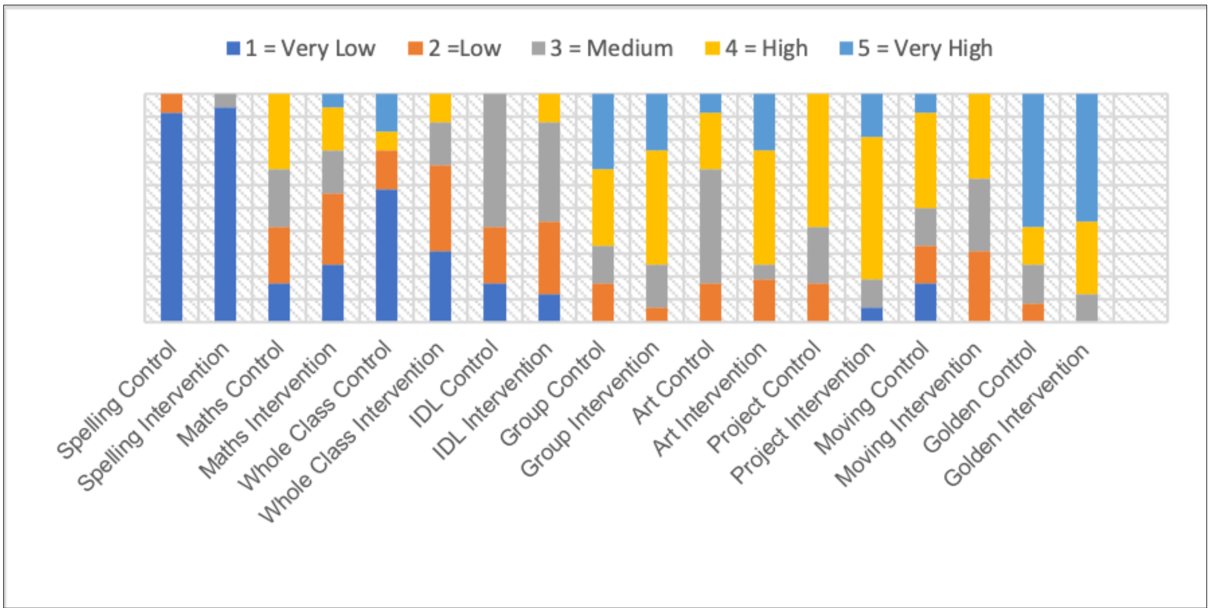


Figure 33: Noise levels by subject and learning activity as observed by the class teachers in the control and intervention classrooms.

5.9.4 Impact of the dynamic soundfield

The teachers in the intervention classroom were asked to evaluate the use of the dynamic soundfield system at the end of the study, the results are presented in Figure 34. The first two questions were about whether the learners in the classroom could hear and respond better to the teacher. In keeping with the findings from the logistic regression model, there was a positive response to hearing and responding to the class teacher. Eighty-one per cent of teachers felt that the learners responded better to spoken instructions with 88 per cent feeling that there was improved access to hearing the teacher. There was no strong agreement on whether the learner's attention span improved after having access to the dynamic soundfield system. Fifty-six per cent (n=9) of teachers agreed with 44 per cent either unsure or in disagreement. This is consistent with the learner's survey where a higher number of learners in the intervention classroom reported improved focus, but this was not significantly different. The class teachers did not feel that there was a change in the levels of background noise during the longitudinal study which is again consistent with the findings from the learner questionnaire.

The teachers were asked if they needed to raise their voices as often. This question was included as it has been suggested that a dynamic soundfield system can reduce vocal strain (da Cruz et al., 2016). Once again there was not a strong response with just over 50 percent providing a positive answer. Furthermore, 69 percent of teachers did not feel self-conscious using the system. As discussed in Chapter 4, there was a high uptake in the use of the system throughout the study.

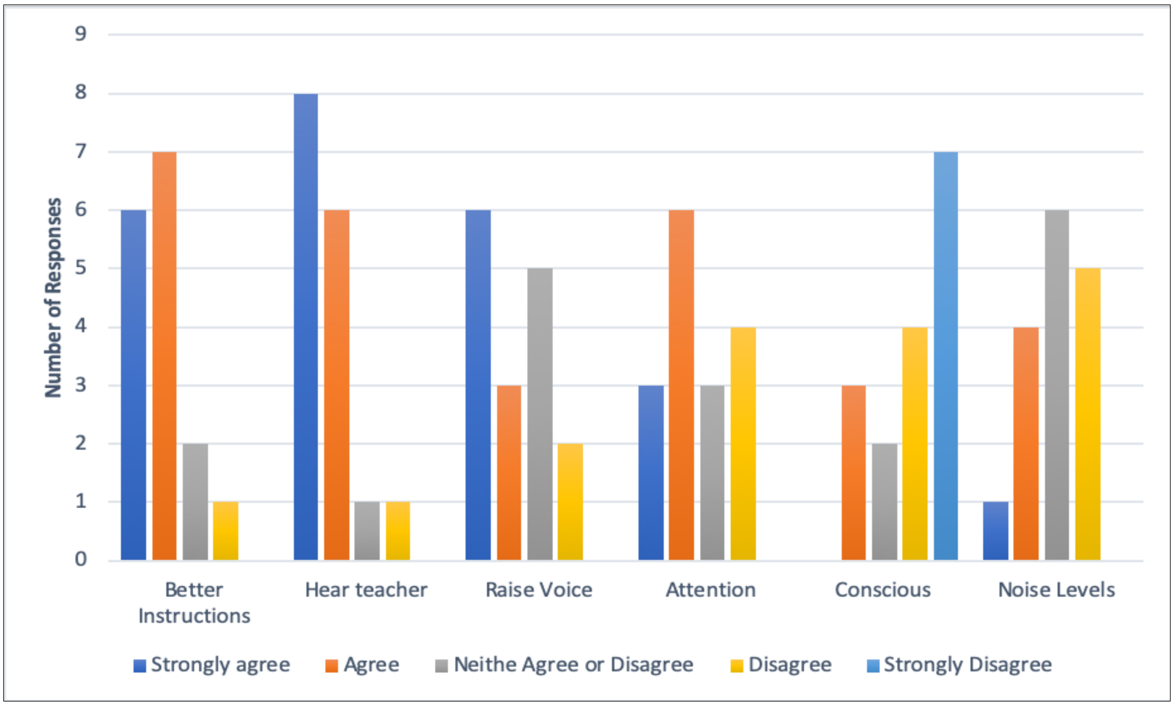


Figure 34: Findings from the teacher questionnaire on the impact of the dynamic soundfield system.

5.10 Discussion of noise levels in the classroom

This research completed the first occupied noise surveys in Scottish schools during lessons in IDL, literacy, and numeracy since the introduction of the CfE. The results show that noise levels were generally higher in mathematics than in literacy lessons. IDL had the highest recorded noise levels which is unsurprising as this is the curricular area more commonly associated with active learning. Previous studies carried out in urban schools by Shield and Dockrell (2008) found that mathematics and science scores were the most significantly affected by background and underlying noise levels. Interestingly, the largest attainment gap in Scottish schools is observed in numeracy (Scottish Government, 2015b).

The overall mean noise level across all three subjects was 64.2dB(A) L_{Aeq} in the control and 64.53dB(A) L_{Aeq} in the intervention classrooms. The results are in keeping with previous studies in both the primary and secondary sectors. Moodley (1989) observed average noise levels of 65dB (A) in both primary and secondary classrooms. Shield and Dockrell (2004) also observed mean values of 68.9dB (A) L_{Aeq} for Year 3 classrooms (age range 7-8 years old) in the 142 London schools surveyed. A survey of 274 lessons in the secondary sector recorded overall noise levels at 64.2dB(A) L_{Aeq} (Shield et al., 2015).

The findings from the noise questionnaires are also corroborated by previous studies. Surveys in urban schools found that noise from young people and adults in corridors and noise from other classrooms was an audible source of internal noise pollution (Shield and Dockrell, 2004). Approximately 79 per cent of learners identified noise from other classrooms and movement around the school as the most commonly observed noise sources that were external to the classroom but internal to the school. This study was unique as it investigated the effects of noise on social deprivation using a large sample size. The results showed that learners from SIMD 1 in non-amplified classrooms were more aware of the noise generated by learners in other classrooms or from people moving around the school than SIMD 5 learners. This is a new finding, not recorded in any published study before.

These findings carry implications for schools. The analysis of the teacher's timetables discussed in Chapter 4 showed that young learners spend over 10 per cent of their week in curricular activities outside the classroom. The curricular areas identified with moving between the classroom and other parts of the school include Physical Education, Music, Drama, ICT and Assembly. In addition, there is movement at the start of the day, break, lunch and home time. Considering most schools in Scotland have classes across the seven primary stages it can be expected that the noise created by the movement of people is a prevalent intermittent noise source.

The findings from this study also indicate that the most commonly observed internal noise pollution was a result of the 'chatter and clatter' created by the learners in the room. This is consistent with previous studies that have shown that the majority of internal noise pollution is occupant generated (Shield and Dockrell, 2004). The present study also demonstrated that SIMD 1 learners were more affected by noise pollution created by fidgeting and movement than SIMD 5 learners. The findings relating to social deprivation and noise are novel and so this research extends the literature by highlighting the different effect that noise has on SIMD 1 learners for the first time. The current study has also shown that the type of learning activity is closely associated with the noise levels generated. Learning activities that are linked with active learning, a key element of the CfE generated the highest level of noise. This trend aligns with previous research in both primary and secondary schools (Shield and Dockrell, 2008, Shield et al., 2015).

The results relating to the challenges of hearing the teacher easily in noise partially support previous research that suggests that young people are adversely affected by listening in the presence of noise (Johnstone and Litovsky, 2006, Blandy and Lutman, 2005) and so require a larger SNR. Bradley and Sato (2004) found that young learners similar in age to the participants in this research required an SNR of 12.5 dB to achieve speech intelligibility performance on par with adults. Dance et al. (2018) investigated the effectiveness of dynamic soundfield in different acoustic conditions and found the system was most effective when noise levels were above 75dB (A). The average occupied noise levels across all three subjects for learning activity A2, which was the activity described in questions 1 and 2 was 57.9 $_{LAeq.}$ It might be expected that when noise levels are low there would a limited dynamic soundfield advantage, especially in question 1 when the listener has access to visual cues. The findings from the current study suggest that this is not the case. Even at the lower background and underlying noise levels, the learners in the intervention classrooms reported that it was easier to hear the teacher easily compared to the control. The findings from the current study suggest that the difficulties experienced by young learners when accessing speech in a classroom are at levels lower than previously reported.

The findings have implications for educators and policymakers who require to address the negative effects that classroom noise has on young learners. The implications are significant as the results suggest that the adaptive amplification of the teachers' voice provides learners an advantage compared to the control. The enhanced SNR produced by good C_{50} classrooms is not sufficient on its own to provide easy access to speech. As discussed in Chapter 2, young people do not always have the experience and knowledge to effectively perceive speech in a multi-talker environment. Differences in the speech characteristics of the talker such as a male target and female distractor are thought to reduce the effects of masking (Cherry, 1953, Freyman et al., 2004). The results from the questionnaire would suggest that dynamic soundfield contributes towards the release from masking. It is unclear whether this is due to the improved SNR, the amplification of the teacher's voice making it distinct from the unamplified voices in the room or both.

The classroom scenario with the largest odds ratio was question two when the learner could not see the face of the teacher. Although noise levels were described as low, there were limited visual cues available to the listener. Listening in noise questionnaires are generally used with deaf learners to identify the learning and teaching experiences that are most challenging (Anderson, 1998). The present study also suggests that young learners with normal hearing thresholds struggle to listen easily in noise and so schools should consider gathering their views on listening in the classroom. Furthermore, organisations such as the National Deaf Children's Society recommend good practice guidelines and teaching strategies to promote effective listening in the classroom. These include closing doors and windows, encouraging a quiet working environment, the class teacher standing at the front of the room facing the class when teaching and the use of soundfield to amplify the teachers' voice (NDCS, 2001). The results from the questionnaire would suggest those teaching strategies usually associated with deaf inclusion would be beneficial to young learners with typical hearing thresholds.

5.11 Key findings

The first purpose of this chapter was to explore the noise levels in the intervention and control classrooms to determine if there was a significant difference. The noise level in the classroom was a situational variable that could compromise the outcome of this study. The second purpose of the chapter was to measure the impact that dynamic soundfield had on hearing the teacher easily through the use of a learner and teacher questionnaire. The main outcomes are as follows:

- In relation to the secondary research question 1, there was not a significant difference in the noise levels observed in numeracy, literacy, and IDL in the control and intervention classrooms. Noise levels were higher in the curricular areas of IDL and numeracy than literacy.
- Noise levels vary depending on the learning activity. Noise levels were higher for activities associated with moving and talking. The highest noise levels recorded were for group work.
- Learners from SIMD 1 were more affected by the noise generated by people moving around the school or other classrooms than learners from SIMD 5. This significantly affected learners in the control classrooms only.
- Learners from the most deprived areas in both the control and intervention classrooms were more aware of noise generated by people retrieving things from bags and scrapping of chairs than learners from the least deprived areas.

- Overall, learners from both the control and intervention classrooms had a negative association with noise levels in the classroom.
- In relation to secondary research question 4, dynamic soundfield technology significantly improved the ability of young people to hear the teacher easily in noise in all six different listening scenarios. This was whether noise levels were low or high.
- The listening scenario with the largest odds ratio was when noise levels were low but there were no visual cues available to the listener. This suggests that having visual cues supports the listening experience in the classroom and that the amplified voice of the teacher helps with the release from masking.

The data collected from the classroom surveys indicate that noise levels were not a confounding variable in this research. Furthermore, the results from the binary logistic regression model also demonstrate that dynamic soundfield technology is a significant predictor for hearing the teacher well. The next chapter will examine the results of the pre and post Achievement for Excellence assessments.

Chapter 6
Results - AfE (InCAS)

6.1 Aims of the chapter

This chapter will address primary research aims one and five of this thesis (see section 3.2.1) by analysing the pre-test and post-test AfE (InCAS) modules and subtest results using a two-way mixed model ANOVA to identify any significant main effects and interactions. An exploration of the relationship between the acoustic environment and dynamic soundfield on each dependent variable will also be undertaken. The chapter will begin by discussing the data collection methods, statistical analysis procedures, effect size and the process used to ensure the data conforms to the assumptions of parametric testing. Thereafter, a comparison of the mean scores from the PIPS assessments (Early Reading, Early Mathematics and Phonological Awareness) completed by the participants at the start and end of the first year in school is undertaken, to identify any broad academic differences in the control and intervention classes. The chapter will conclude by discussing the key findings.

6.2 Data collection and analysis

6.2.1 Data collection

The individual participant AfE (InCAS) data was collected from the transfer files provided by the CEM at the University of Durham and shared with the individual schools in Fife. Each school has a CEM account and the Primary 3 records were harvested at the pre and post-intervention stage. The raw assessment files provide the Scottish Candidate Number as well as information on gender and date of birth. The Scottish Candidate Number was used to harvest the Fife Council's data management system for each of the PIPS assessments from Primary 1. The PIPS and AfE (InCAS) data were provided in a .csv file format which was then imported into SPSS version 23 for Mac. All the raw AfE (InCAS) transfer files have the age at test, age equivalent score and age difference score formatted in decimal. There was no data on the acoustic properties of the classrooms the learners attended in primary 1.

Four hundred and ninety-five pre-tests and post-test AfE (InCAS) files were retrieved from the CEM accounts in each school and analysed. Four hundred and eighty-nine primary one PIPS records were extracted from the Council's data management system, which is an attrition rate of 1.21 per cent. There was an even split in the attrition figures: three students each in the control and intervention groups. In the control condition two learners were from SIMD 5 quintile, and one from SIMD 1. In the intervention classrooms, all the attrition was from SIMD 1. The reason for the missing data is unknown. The learners may not have been educated in Fife at the time or were absent during the assessment period. PIPS scores have a normal distribution with a mean of fifty and a standard deviation of ten. To convert the PIPS scores to an AfE (InCAS) data format the standardised scores are arithmetically multiplied by 1.5 and then twenty-five is added. All scores were converted using this formula. The AfE (InCAS) assessments have a mean of 100 and a standard deviation of fifteen. 68% of all learners will have a score between 85 and 115.

6.2.2 Data analysis

Two methods often used to analyse the outcomes of a pre-intervention and post-intervention design are an analysis of covariance (ANCOVA) and ANOVA. Jamieson (2004) cautions against the use of an ANCOVA that uses the post-test score as the dependent variable and pre-test as the covariate or a modified version in which the difference score (pre-test minus post-test) is used as the dependent variable and the pre-test as the covariate for non-randomised groups. Both approaches produce biased conclusions due to measurement error when comparing naturally occurring groups. Although allocation to the control and intervention classrooms in this research was achieved through concealed randomisation it was not through the classic random assignment and so ANCOVA may produce biased outcomes.

A two-way Mixed ANOVA of change was used to explore how the dynamic soundfield intervention affected AfE (InCAS) outcomes compared to the control classes. Cribbie and Jamieson (2004) observe that a pre-test and post-test study design is still commonly used to measure changes over time in two or more groups, where one is exposed to an intervention. There are two advantages over a post-test only design: the pre-test stage provides data on individual differences which reduces error variance and so increases power. Furthermore, differences between the groups at the pre-test stage can be factored into the analysis. A two-way mixed ANOVA compares the mean difference between groups based on a within-subject factor (time) and a between-subject factor (condition) (Leech, 2005). One of the strengths of a mixed ANOVA is the ability to test the combined effects or interaction of two independent variables on a dependent variable (Field, 2013). This research has two independent factors, one within-subject factor, with two levels (pre-test, time point 1 and post-test, time point 2) and one between-subject factor, with two levels (control and intervention). The dependent variable was the pre-test and post-test scores from the individual modules and subtests that form part of the AfE (InCAS) suite of assessments. The interaction was the effect of condition x time on the dependent variable.

Using Post-hoc tests to follow up any significant interactions was not applicable as each of the between and within-subject factors only had two levels. Instead, the strength of any interaction was measured using independent (between-subjects) and paired (within-subjects) *t*-tests. When running several *t*-tests simultaneously there is an increased risk of obtaining a Type I error and so a Bonferroni correction was applied. A correction is achieved by dividing the number of comparisons being made with the α value (0.05).

Finally, as revealed in Chapter 4, initial analysis using Chi-Square and independent *t*-tests were conducted on the characteristics of the control and intervention groups and found there was no significant difference between each group on gender, ethnicity, ASN, age, and attendance.

6.2.3 Effect size

In a two-way Mixed ANOVA, the degree of association between the main effect or interaction and the dependent variable is measured using partial Eta squared (η_p^2). Partial Eta squared is the ratio of variance accounted for by an effect, plus the effect and its associated error. The following formula applies where SS refers to the sum of square (Richardson, 2011):

$$\eta_p^2 = \frac{SS\ (Between\ Groups)}{[SS(Between\ groups) + SS(Within\ groups)]}$$

Turturean (2015) describes the effect size as a name given to indices that measure the magnitude of a treatment. The following guide was recommended by Cohen (1988) when measuring the effect size of η_p^2:

Small	0.01
Medium	0.06
Large	0.14

The strength of the interactions was measured using Cohen's $\boldsymbol{d}$ effect size which was calculated using the following formula where m_2 and m_1 are the means of the control and intervention groups and s_{pooled} is the pooled standard deviation (Nakagawa and Cuthill, 2007):

$$d=\frac{m_2-m_1}{s_{pooled}}$$

The following guide was used to measure Cohen's $\boldsymbol{d}$ values (Turturean, 2015):

Small	$\boldsymbol{d} \leq 0.2$
Medium	$0.2 < \boldsymbol{d} \leq 0.5$
Large	$0.5 < \boldsymbol{d}$

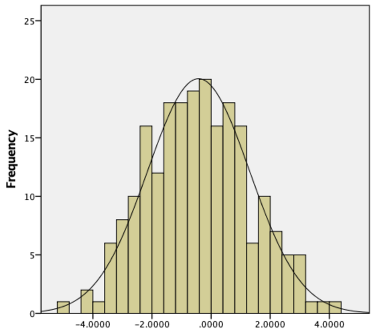
6.3 Analysis of assumptions

One of the assumptions of a two-way Mixed ANOVA is the dependent variables are normally distributed. To explore the assumption of normality Histograms, Q-Q-plots and calculation of skewness and kurtosis values were assessed. Another assumption is that variance is equally spread between the control and intervention groups. The homogeneity of variance was assessed using the Levene's test and the Hartley's F_{max} test.

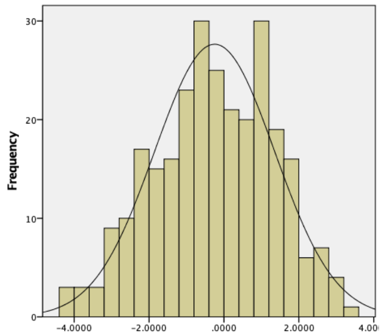
6.3.1 Distribution of the data

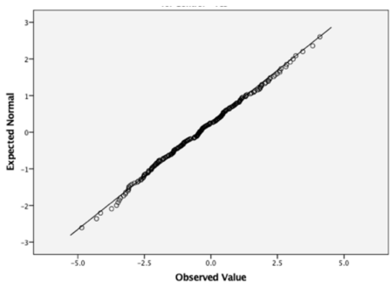
For large sample sizes (>300) the testing of normal distribution relies on histograms and the absolute values of skewness and kurtosis, without reference to z-values. Furthermore, the use of the Shapiro-Wilk normality test is generally not recommended for larger sample sizes as small deviations from normality can produce significant results and the assumption of normality may be incorrectly rejected (Field, 2013). Kurtosis values of >7.0 and skewness values of ≤2.0 and ≥-2.0 were used to reference substantial non-normality (Hae-Young, 2013). The histograms and Q-Q-plots for the Developed Ability AfE (InCAS) module is presented in Figure 35. Appendix 16 provides the histograms and Q-Q-plots for all AfE modules and subtests by control and intervention groups. As can be seen inspection of the histograms in Figure 35, in terms of the Developed Ability outcome measure the distribution of the data is normal. This suggests that the assumption of normality is not violated. This was replicated in all modules and subtests.

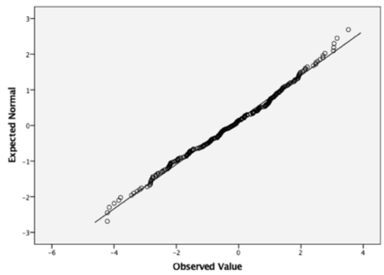
Pre-Intervention Developed Ability Control



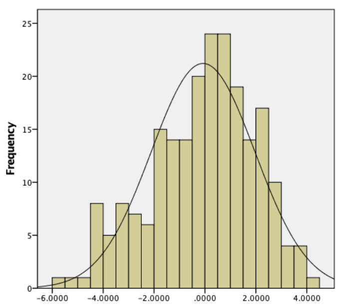
Pre-Intervention Developed Ability Intervention



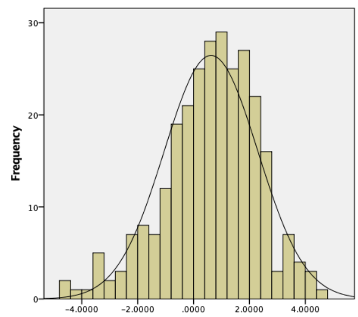


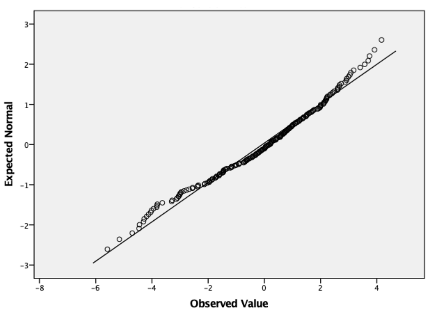


Post-Intervention Developed Ability Control



Post-Intervention Developed Ability Intervention





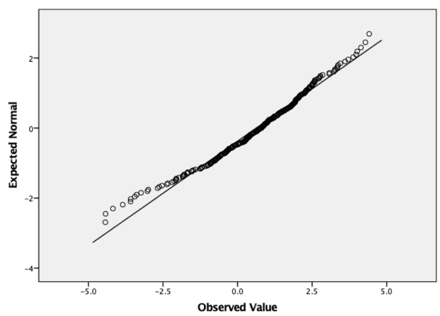


Figure 35: Histograms, Q-Q-plots, for the Developed Ability module in the control and intervention classrooms.

Table 44 presents the skewness and kurtosis values for each of the AfE (InCAS) outcome measures by the control and intervention group. Once again, all are below the recommended absolute values and so the assumption that the data is normally distributed has not been violated.

	Control		Intervention	
	Skewness	Kurtosis	Skewness	Kurtosis
Developed Ability (Pre)	0.103	-.0372	-0.164	-0.470
Developed Ability (Post)	-0.407	-0.351	-0.505	0.355
Picture Vocabulary (Pre)	0.038	-0.252	-0.274	-0.322
Picture Vocabulary (Post)	-0.371	-0.290	-0.303	-0.187
Non-verbal (Pre)	0.262	-0.614	-0.067	-0.997
Non-verbal (Post)	-0.360	-0.372	-0.561	0.412
Reading (Pre)	0.675	-0.236	0.273	-0.663
Reading (Post)	0.005	-0.740	-0.035	-0.466
Word Recognition (Pre)	0.575	-0.115	0.159	-0.683
Word Recognition (Post)	-0.128	-0.871	-0.131	-0.465
Word Decoding (Pre)	0.661	-0.224	0.349	-0.853
Word Decoding (Post)	0.140	-0.604	-0.108	-0.632
Comprehension (Pre)	0.494	-0.443	0.282	-0.699
Comprehension (Post)	-0.029	-0.694	-0.022	-0.407
Spelling (Pre)	0.360	-0.180	0.308	0.161
Spelling (Post)	-0.214	-0.544	-0.268	-0.225
Mental Arithmetic (Pre)	-0.071	0.277	-0.519	0.011
Mental Arithmetic (Post)	-0.579	-0.178	-0.742	0.801
General Maths (Pre)	-0.385	1.151	-0.210	1.377
General Maths (Post)	-0.106	0.012	0.042	0.994
Number 1 (Pre)	-0.326	0.453	-0.575	2.057
Number 1 (Post)	-0.328	0.076	-0.100	1.321
Number 2 (Pre)	-0.005	0.387	-0.412	1.482
Number 2 (Post)	0.101	0.699	0.175	0.539
Data (Pre)	-0.367	0.296	-0.069	-0.254
Data (Post)	0.105	0.328	-0.352	0.518
MSS (Pre)	-0.133	0.396	-0.018	0.366
MSS (Post)	0.008	0.111	0.340	-0.420

Table 44: Skewness and kurtosis values for both the control and intervention classrooms.

6.3.2 Homogeneity of variance

To test that variance was equally distributed across both the control and intervention groups, each of the fourteen AfE (InCAS) modules and subtests at the pre-intervention and post-intervention stage were analysed using the Levene's test. Variance is assumed to be equal if the results are non-significant ($p>0.05$). At the pre-intervention stage, Non-Verbal Ability ($p=0.044$) and Word Decoding ($p=0.007$) violated this assumption. At the post-intervention stage Word Recognition ($p<0.001$), Word Decoding ($p=0.011$), Comprehension ($p=0.004$), Spelling ($p=0.038$) Standardised Reading ($p=0.002$), Non-Verbal Ability ($p<0.001$), Standardised Developed Ability ($p=0.006$) and Mental Arithmetic ($p=0.003$) violated the assumption of homogeneity. All of the other AfE (InCAS) modules and subtests were non-significant.

Field (2013) observes that for large sample sizes, Levene's test can be too sensitive to small differences in variance and so produce a false-positive result. It is recommended that Hartley's F_{max} test should be used as a cross-reference check, and this is calculated by dividing the largest variance with the smallest variance of each AfE (InCAS) modules and subtests. If the figure is close to one, then homogeneity of variance can be assumed. Where this condition is not met then the number of groups (K) and the degrees of freedom (number of participants in a group minus one) is compared against Hartley's F_{max} Table to obtain a value. If the calculated variance value is smaller than the table value, then homogeneity is assumed.

The Hartley's F_{max} values for pre-intervention scores are close to one and so homogeneity is assumed: Non-Verbal Ability ($F_{max}=1.19$) and Word Decoding ($F_{max}=1.25$). The post-intervention data produced a similar outcome: Word Recognition ($F_{max}=1.44$), Word Decoding ($F_{max}=1.29$), Comprehension ($F_{max}=1.37$), Spelling (F_{max}=1.35) Standardised Reading (F_{max}=1.41), Non-Verbal Ability (F_{max}=1.36), Standardised Developed Ability (F_{max}=1.36) and Mental Arithmetic (F_{max}=1.38).

6.4 Primary 1 PIPS standardised assessments

To determine if there was a significant difference between the control and intervention classes in early mathematical ability independent *t*-tests ($\alpha=0.05$) were performed on the baseline and follow-up scores. The results suggest that when entering school there was a significant difference between the learners in the control and intervention classrooms, $t(487)=-2.14$, p=0.033. Inspection of the means shows that the control classrooms (M=96.62, SD=17) were significantly behind the intervention classrooms (M=99.87, SD=16). By the end of the first year in school, both groups demonstrated improvement: control (M=101.31., SD=17) and intervention (M=104.78, SD=15). The difference between the groups was also significant, $t(487)=-2.34$, p=0.020). These overall results suggest that there were some significant differences in early mathematical ability observed during the first year in school.

In early reading at baseline, both the control (M=94.62, SD=16) and intervention (M=98.07, SD=16) groups recorded marginally lower mean scores than those observed in the early mathematics module. The analysis revealed a significant difference between the groups when starting school, $t(487)=-2.33$, p=0.020. Unsurprisingly, by the end of the school year, both groups demonstrated an overall improvement in mean scores: control (M=101.34, SD=17) and intervention (M=103.82, SD=16). Furthermore, the difference between the two groups at the end of the year was non-significant, $t(487)=-1.13$, p=0.260. The results suggest that both groups made progress during the first year in school with the learners allocated to the control classrooms making the largest amount of gain.

In phonological awareness, there was a significant difference between the control (M=95.08, SD=15) and intervention (M=99.02, SD=15) groups at the start of primary one, $t(487)=-2.90$, $p=0.004$. By the end of primary one, the control group (M=97.87, SD=13) increased their mean score and the intervention group (M=99.8, SD=13) made minimal progress. At the end of the first year of school, the gap between the two groups reduced. Once again, the control learners demonstrated the greatest amount of improvement. There was also no significant difference between the control and intervention classrooms by the end of primary one, $t(487)=-1.73$, $p=0.084$.

6.4.1 Summary Primary 1 PIPS standardised assessments

As discussed in Chapter 3, the allocation of participants to the control and intervention groups was by concealed randomisation. To determine if there was an inherent bias between the learners in the control and intervention classrooms the standardised assessments completed by the participants during their first year in school were examined. These were the only standardised assessment completed by the learners and graded by an external body before the current research at the start of primary 3. Four hundred and eight-nine records were examined. In all modules, there was a significant difference between the learners in the control and intervention classrooms when entering compulsory education. By the end of the primary 1 year, there was no significant difference between the control and intervention classrooms in reading and phonological awareness. The control classrooms demonstrated the largest amount of improvement. Early mathematics demonstrated a significant difference between the classes at the end of the first year in school. The intervention classrooms had a higher mean score at the end of primary one than the control.

6.5 AfE (InCAS) assessments -Developed Ability

A series of two-way mixed ANOVAs were conducted with time and condition as the independent factors and the pre and post-test age difference scores from the Developed Ability subtests (age at assessment minus age equivalent score) as the dependent variable. For the overall Developed Ability module, the dependent variable is the pre-test and post-test age-standardised scores which are calculated using the Picture Vocabulary and Non-Verbal Ability subtest scores using the formula presented in section 3.10.1. The results from the mixed ANOVAs are presented in Table 45

	Within-Subject			**Between-Subject**			**Interactions**		
α=0.05 (1,493)	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
Picture Vocabulary	39.67	0.001	0.074	5.75	0.017	0.012	16.64	0.001	0.033
Non-Verbal Ability	100.02	0.001	0.169	4.24	0.040	0.009	5.39	0.021	0.011
Developed Ability	113.79	0.001	0.188	7.16	0.008	0.014	17.15	0.001	0.034

Table 45: Results from the two-way mixed ANOVAs conducted on the Developed Ability module and subtests.

6.5.1 Picture Vocabulary subtest

As Table 45 illustrates, the change in scores between the pre-test and post-test scores was significant. The magnitude of the difference between the two-time points was medium. The two-way mixed ANOVA also revealed a significant main effect of condition. Importantly, a significant interaction was observed between condition and time, this suggests that the effect of one independent variable on the dependent variable depends on the level of the other independent variable. The effect size for the interaction was small. Heiman (2013) observes that it is generally advisable to use interaction plots to support the interpretation of any significant interaction. Figure 36 illustrates that the change in pre-test and post-test scores are greater in the intervention classrooms (green line) compared to the control (blue line). The non-parallel configuration of the two lines signifies the significance of the interaction.

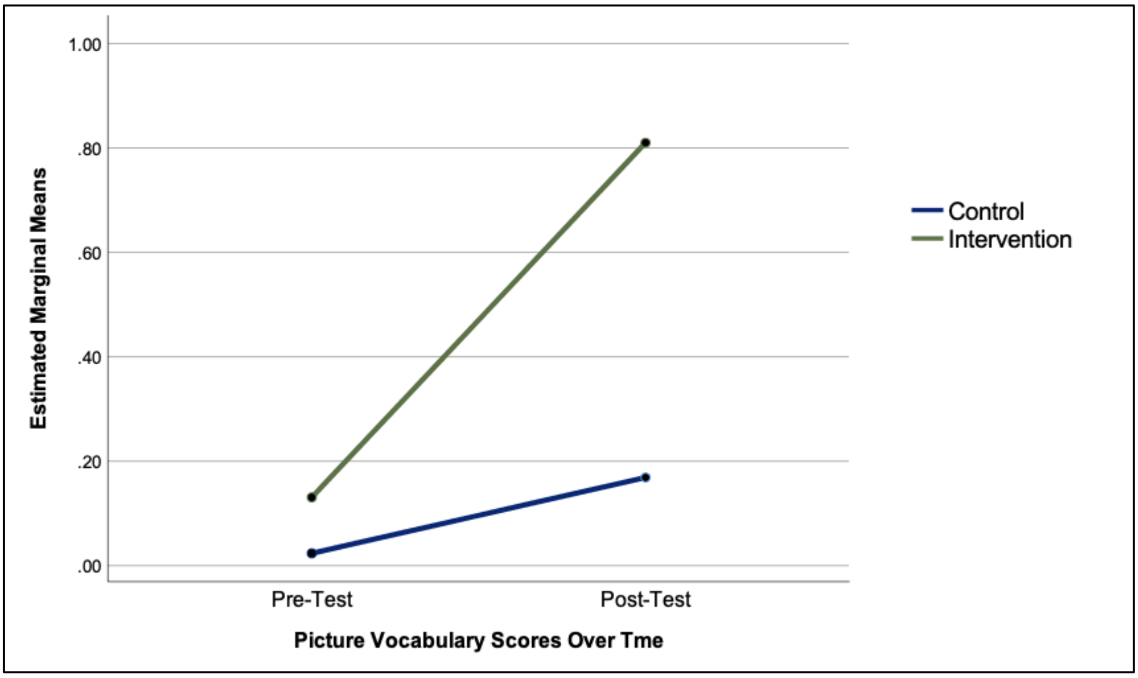


Figure 36: Interactional line chart showing the change in picture vocabulary scores from the pre-test and post-test assessments in the control (blue line) and intervention (green line) classrooms.

As the interaction between condition and time was significant, it is recommended to ignore the two main effects and instead examine the differences between and within the control and intervention classrooms for each of the pre and post-test scores (Field, 2013). To test the significance of the interaction two paired *t*-tests (two-tailed) and two independent *t*-tests (two-tailed) for each level of the independent variables were performed. To protect against a Type I error a Bonferroni correction was applied ($\alpha/4$ =0.0125). The results show that there was no significant difference between the pre-test and post-test scores in the control classrooms, $t(216)=-1.439$, p=0.124. In contrast, there was a significant difference in the intervention classrooms, $t(277)=-8.02$, $p<0.001$. The results indicate that learners in classrooms exposed to dynamic soundfield significantly improved their knowledge of picture vocabulary over the academic year compared to the control classrooms. Between-subject, the analysis showed that before the installation of the dynamic soundfield there was a non-significant difference between the two groups, $t(493)=-0.658$, p=0.511. The post-test analysis demonstrated a significant difference between the two groups, $t(493)=-3.61$, $p<0.001$. The magnitude of the difference was medium ($\boldsymbol{d}$=0.33). Both these factors, collectively, drove the significant interaction effect.

6.5.2 Non-Verbal Ability subtest

A two-way mixed ANOVA showed there was a significant main effect of time. Inspection of the means reveals, at the pre-intervention stage both the control (M=-1.04, SD=2.1) and intervention (M=-0.876, SD=2.3) classrooms were behind their expected age equivalent level. By the end of the study, the control group was approximately three months behind their age equivalent level and the intervention group was three months ahead. This generated a large effect size. The analysis also showed a significant main effect of condition, although the effect size was small. As Table 45 illustrates, the condition x time interaction was also significant, indicating that the change in the pre and post-test scores were dependent on the classroom the learners attended. Further analysis was performed to decompose the interaction and establish the factors driving it.

The analysis demonstrated there was a significant difference between the pre-test and post-test scores in both the control, $t(216)=-5.42$, $p<0.001$ and intervention classrooms, $t(277)=-8.94$, $p<0.001$. The change in scores was greater in the intervention group (M=1.15, SD=2.1) compared to the control (M=.719, SD=1.9). The pre-intervention analysis showed that there was no significant difference between the groups, $t(493)=-0.816$, $p=0.415$. Interestingly, when the post-trial analysis was undertaken a significant difference between the control and treatment groups was revealed, $t(493)=-2.83$, $p=0.005$. The magnitude of the difference between the groups was small to medium ($\boldsymbol{d}$=0.26). The learners in the dynamic soundfield classrooms improved to a greater extent in non-verbal ability compared to the control classroom and this resulted in a significant difference between the groups at the end of the study. Both these results produced a significant interaction effect, and this is graphically illustrated in Figure 37. Overall, the results suggest that the improved scores in the intervention classrooms might be attributable to the beneficial effect of the dynamic soundfield system.

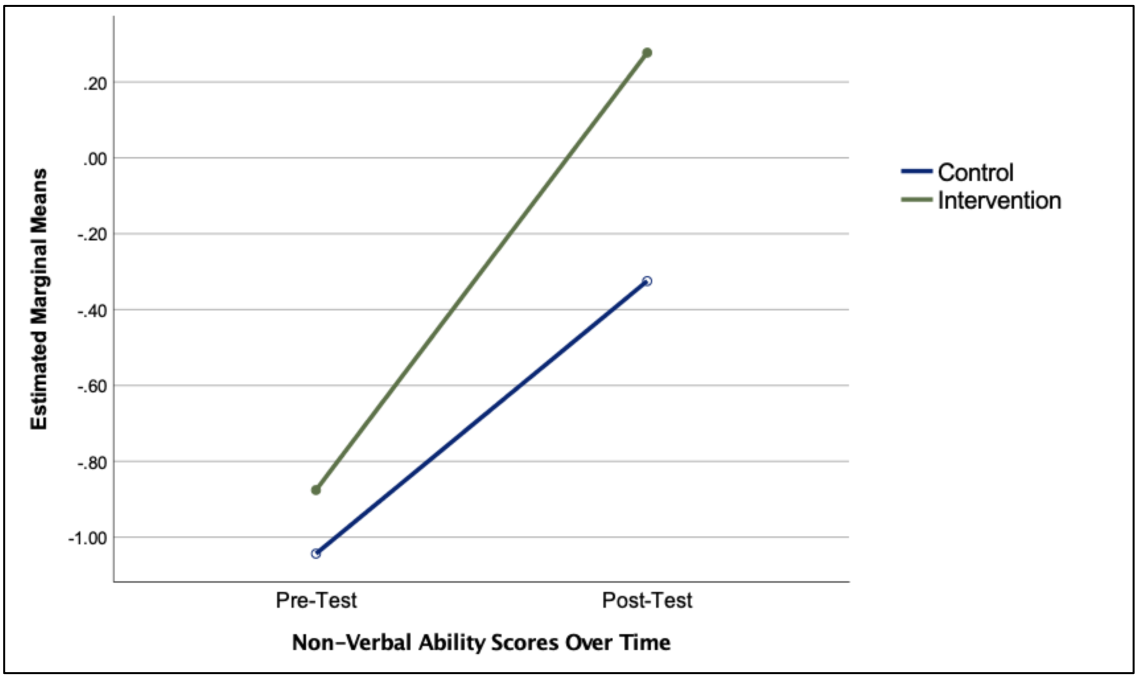


Figure 37: Interactional line chart showing the change in non-verbal ability scores from the pre-test and post-test assessments in the control (blue line) and intervention (green line) classrooms.

6.5.3 Developed Ability module

As Table 45 illustrates the two-way mixed ANOVA on the standardised Developed Ability module showed a significant main effect of time, condition and condition x time interaction. The significant interactions are illustrated in Figure 38, which shows that the intervention classrooms (green line) increased their scores from baseline to a greater extent than the control (blue line). Further analysis indicated there was a significant difference in the pre-test and post-test scores in both the control, $t(216)=-4.34$, $p<0.001$ and intervention groups, $t(277)=-11.21$, $p<0.001$. Inspection of the means confirms that the increase was higher in the intervention classrooms (M=-7.77, SD=11.6) compared to the control (M=-4.45, SD=12.34).

Figure 38 also shows the difference between the classrooms exposed to dynamic soundfield and control appears more marked at the end of the intervention. At the pre-trial stage, there was a non-significant difference between the two groups, $t(493)=-1.03$, $p=0.304$. The standardised scores at the pre-test stage for the control group were 97.07, SD=15 and the intervention group 98.47, SD=15. The post-intervention analysis demonstrated a significant difference between the control and intervention classrooms, $t(493)=-3.82$, $p<0.001$. The standardised scores at the post-trial stage for the control group were 100.50, SD=17 and the intervention group 106.26, SD=15. This suggests that learners exposed to dynamic soundfield amplification improved their developed ability skills at a higher rate compared to the control. The between-group effect size at the end of the intervention in favour of the dynamic soundfield classrooms was medium ($\boldsymbol{d}$=0.35).

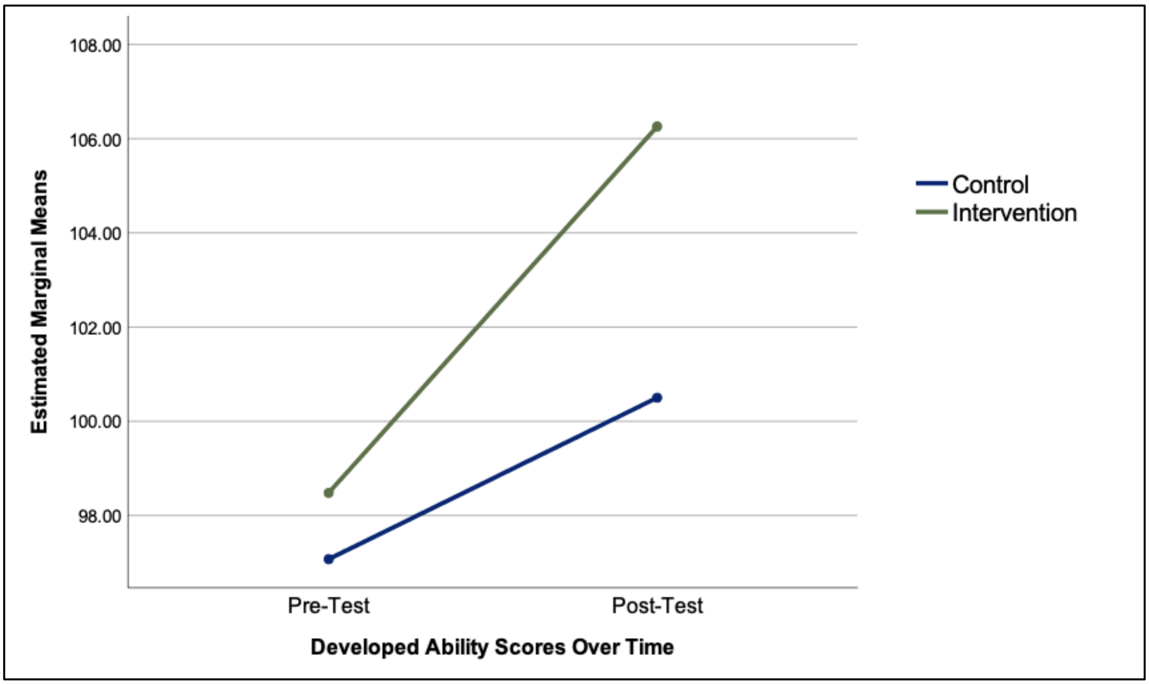


Figure 38: Interactional line chart showing the change in developed ability scores from the pre-test and post-test assessments in the control (blue line) and intervention (green line) classrooms.

6.5.4 Developed Ability Module and Subtests and C_{50}

In Chapter 4, it was established that the research classrooms were categorised as excellent (n=2), good (n=18) and fair (n=5) for speech weighted C_{50}. The two-way mixed ANOVAs were repeated comparing the improvements made in the intervention classrooms on the subtest and standardised scores in the three different acoustic conditions.

In the Picture Vocabulary subtest only classrooms that had good acoustics for speech demonstrated a significant main effect of time, $F(1, 350)=21.32$, $p<0.001$, $\eta_p^2=0.057$; condition, $F(1, 350)=7.24$, $p=0.007$, $\eta_p^2=0.020$; and time x condition interaction, $F(1, 350)=19.60$, $p<0.001$, $\eta_p^2=0.053$. As Figure 39 illustrates the interaction appears to be driven by two contrasting events; the intervention classrooms making approximately eight months of progress from the baseline and the control group exhibiting only modest improvements. Further testing confirmed that the control group made no significant progress from baseline, $t(155)=-0.117$, $p=0.907$. In comparison the intervention classrooms generated a significant improvement, $t(195)=-7.39$, $p<0.001$. Furthermore, before the intervention there were no significant differences between the two groups, $t(350)=-0.850$, $p=0.396$ and at follow-up, the change in scores resulted in a significant difference being observed, $t(350)=-4.00$, $p<0.001$. Examination of these results in conjunction with those discussed in section 6.5.1 suggests that most of the improvements observed in the intervention classrooms were in the classroom environments that were categorised as good for speech clarity, in contrast, the control classroom gains were from a combination of the fair and excellent classrooms. Overall, the results suggest that dynamic soundfield was primarily effective in good acoustic conditions.

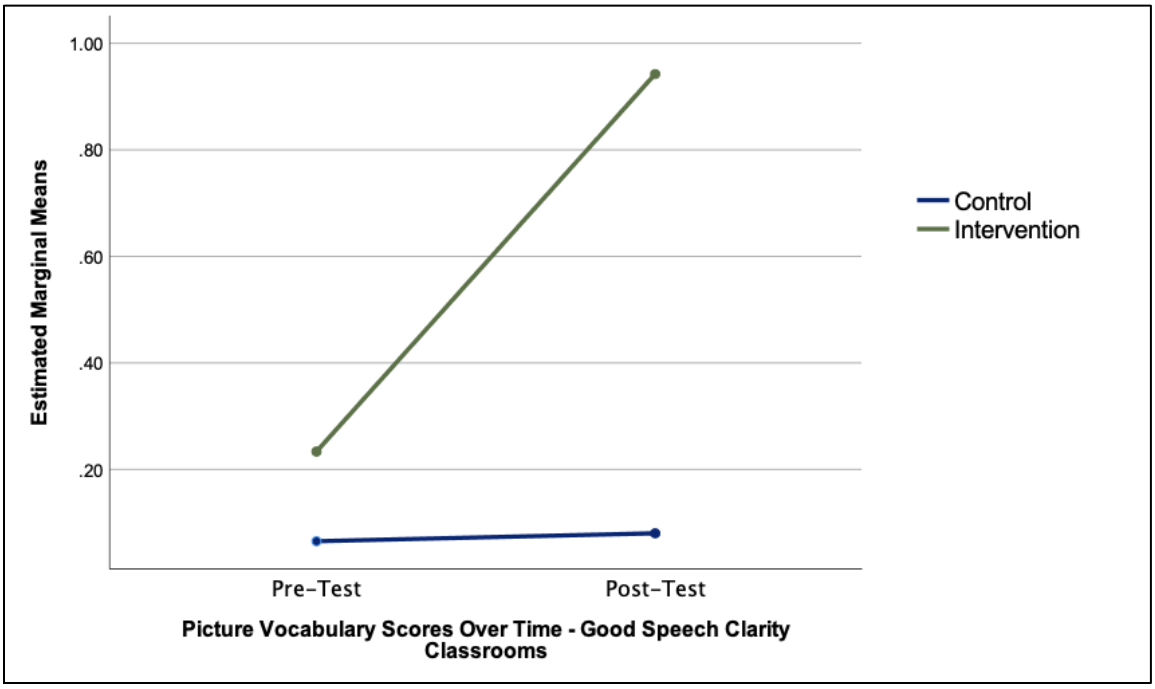


Figure 39: Interactional line chart showing the change in Picture Vocabulary scores from the pre-test and post-test assessments in good speech-weighted C_{50} classrooms (n=18) in the control (blue line) and intervention (green line) classrooms.

Consistent with the results from the Picture Vocabulary subtest, the Non-Verbal Ability test was also primarily effective in the good acoustic classrooms for speech with a significant main effect of time, $F(1, 350)=72.06$, $p<0.001$, $\eta_p^2=0.171$; condition, $F(1, 350)=5.86$, $p=0.016$, $\eta_p^2=0.016$; and interaction, $F(1, 350)=6.21$, $p=0.013$, $\eta_p^2=0.017$. Figure 40 illustrates that the time x condition interaction appears to be driven by the post-intervention scores in the control classrooms being lower than the intervention group. The analysis was run to decompose the interaction further. Both groups demonstrated progress from baseline, with the intervention classrooms (M=1.2, SD=2.2), $t(195)=-7.68$, $p<0.001$ demonstrating the biggest change in score compared to the control (M=0.68, SD=1.9), $t(155)=-4.48$, $p<0.001$. Before the intervention, the difference between the groups was non-significant, $t(350)=-0.965$, $p=0.335$, by the end of the study the differentials between the group was significant, $t(350)=-3.34$, $p<0.001$. This suggests that the classrooms exposed to dynamic soundfield improved their non-verbal ability scores significantly more than the control. This produced a significant interaction effect. Analysis of both the fair and excellent classrooms showed no significant interactions, which would indicate that the improvement in scores between baseline and follow-up may be a result of other variables rather than the intervention. As a result, no further testing was performed.

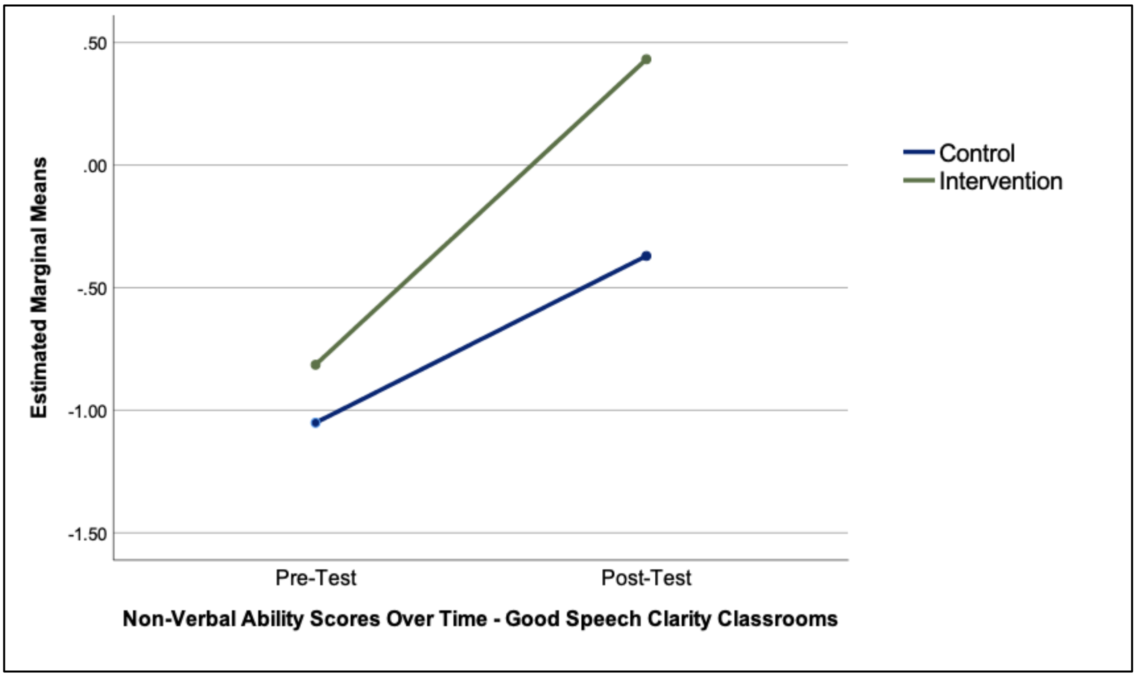


Figure 40: Interactional line chart showing the change in Non-Verbal scores from the pre-test and post-test assessments in good C_{50} classrooms (n=18) in the control (blue line) and intervention (green line) classrooms.

The mixed model ANOVA revealed a significant interaction between the Developed Ability standardised scores over time and the good speech-weighted C_{50} classrooms, $F(1, 350)=19.54$, $p<0.001$, $\eta_p^2=0.053$. A significant main effect, $F(1, 350)=79.16$, $p<0.001$, $\eta_p^2=0.184$ and condition, $F(1, 350)=9.58$, $p=0.002$, $\eta_p^2=0.027$ was also identified. As Figure 41 illustrates the factor that appears to be driving the interaction is the intervention classroom increased its score over time more than the control. Paired *t*-tests confirmed that both classes made significant progress from baseline: control, $t(155)=-3.06$, $p=0.003$ and intervention, $t(195)=-9.86$, $p<0.001$. Unsurprisingly, the change in scores over time was larger in the intervention classrooms (M=8.62, SD=12) compared to the control (M=2.90, SD=12). The independent *t*-test also demonstrated there was no difference between the groups prior to the intervention, $t(350)=-1.26$, $p=0.210$, in contrast at the end of the study the differential change in scores resulted in a significant difference being observed, $t(350)=-4.36$, $p<0.001$. The difference between the groups generated a moderate effect size (d=0.47). Once again, both the fair and excellent conditions showed no significant interactions and so no further testing was performed.

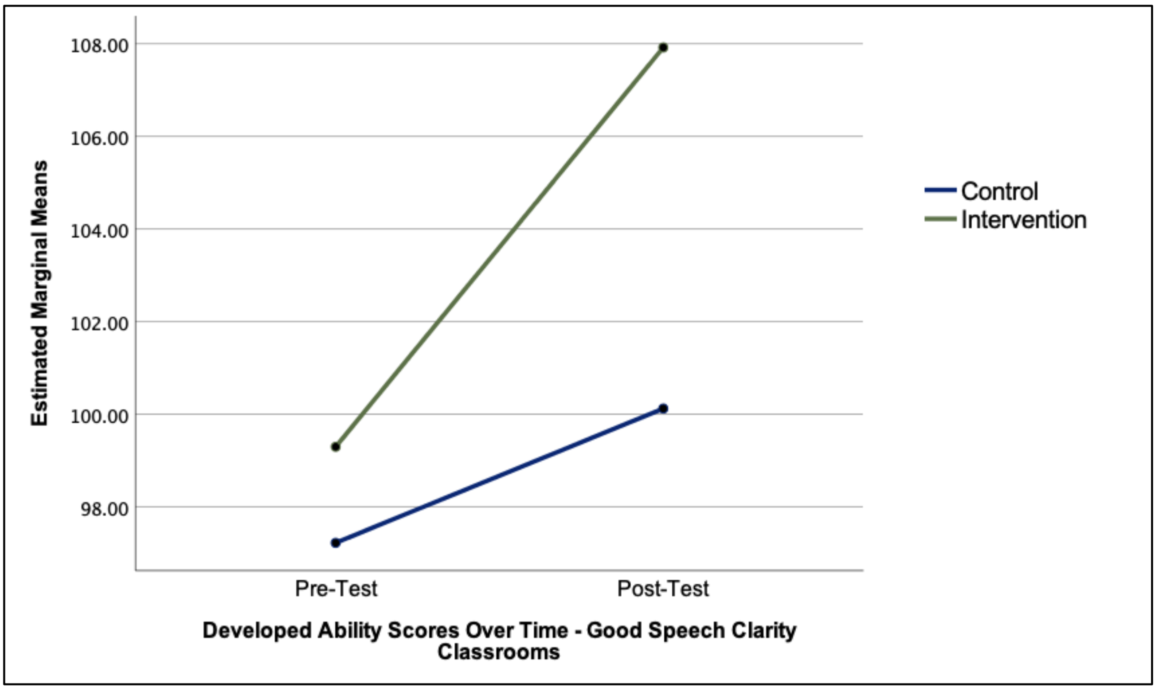


Figure 41: Interactional line chart showing the change in developed ability scores from the pre-test and post-test assessments in good speech-weighted C_{50} classrooms (n=18) in the control (blue line) and intervention (green line) classrooms.

6.5.5 Summary – Developed Ability Module and Subtests

The results show that for all the Developed Ability subtests and standardised module scores, the classrooms fitted with dynamic soundfield achieved an overall higher post-test tariff compared to the control. Further analysis indicated that the dynamic soundfield system was primarily effective in classrooms that were categorised as having good speech-weighted C_{50} values. However, any conclusions about whether the findings can be generalised about the effectiveness of dynamic soundfield and different acoustic conditions should be treated with caution. The learners in the excellent classrooms all come from the least deprived areas of Scotland and the sample size is small. Furthermore, the majority of learners in the fair classrooms are from SIMD 1 and 2.

6.6 AfE assessments – General Mathematics

A series of two (condition) times two (time) mixed ANOVAs examined the effects of time (pre-test and post-test), condition (soundfield and control) and interaction (time x condition) on the General Mathematics module and subtests scores. The subtests were measured using age equivalent scores and the overall General Mathematics module was measured using standardised scores. The results are presented in Table 46.

	Within-Subject			**Between-Subject**			**Interactions**		
α=0.05 (1,493)	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
Number 1	8.76	0.003	0.017	13.26	0.001	0.026	1.27	0.260	0.003
Number 2	39.88	0.001	0.075	11.71	0.001	0.023	4.43	0.036	0.009
Data Handling	34.57	0.001	0.077	14.05	0.001	0.028	4.95	0.027	0.011
MSS	15.10	0.001	0.015	13.66	0.001	0.027	.017	0.896	0.001
General Mathematics	24.22	0.001	0.047	19.30	0.001	0.038	.188	0.644	0.004

Table 46: Results from the two-way mixed ANOVAs conducted on the General Mathematics module and subtests.

6.6.1 Number test 1 subtest

The two-way mixed ANOVA showed there was a significant main effect at the 5% level between the pre-test and post-test scores, and this generated a small effect size. Paired *t*-tests (two-tailed) illustrate that there was no significant difference in age difference scores between time points 1 and 2 in the control classrooms, $t(216)=-1.17$, $p=0.244$. In contrast, there was a significant difference in the classrooms exposed to the dynamic soundfield intervention, $t(277)=-3.20$, $p=0.002$. As Table 46 illustrates, there was a significant main effect of condition, which once again had a small effect size. The analysis also showed that there was a significant difference between the two groups at the pre-intervention, $t(493)=-2.03$, $p=0.043$ and follow-up stages $t(493)=-3.41$, $p<0.001$. There was no significant condition x time interaction, and so the changes in scores may be attributable to other factors. No further testing was performed.

6.6.2 Number test 2 subtest

Number 2 subtest data had a significant effect of time, which generated a medium effect size. Regarding the between-subject factor, a significant main effect was also observed. More importantly, as Table 46 illustrates there was also a significant interaction, indicating that the change in pre-test/post-test scores was dependent upon which classroom the learner attended. The interactions are graphically illustrated in Figure 42, which shows that the scores in the dynamic soundfield classrooms (green line) significantly increased over time more sharply than the control (blue line). Inspection of the means confirms that the change in scores from baseline to follow-up was greater in the intervention classrooms ($M=-0.471$, $SD=1.28$) compared to the control ($M=-0.235$, $SD=1.2$). Interestingly, both the control, $t(216)=-2.95$, $p=0.002$ and intervention classrooms, $t(277)=-6.14$, $p<0.001$ demonstrated significant progress.

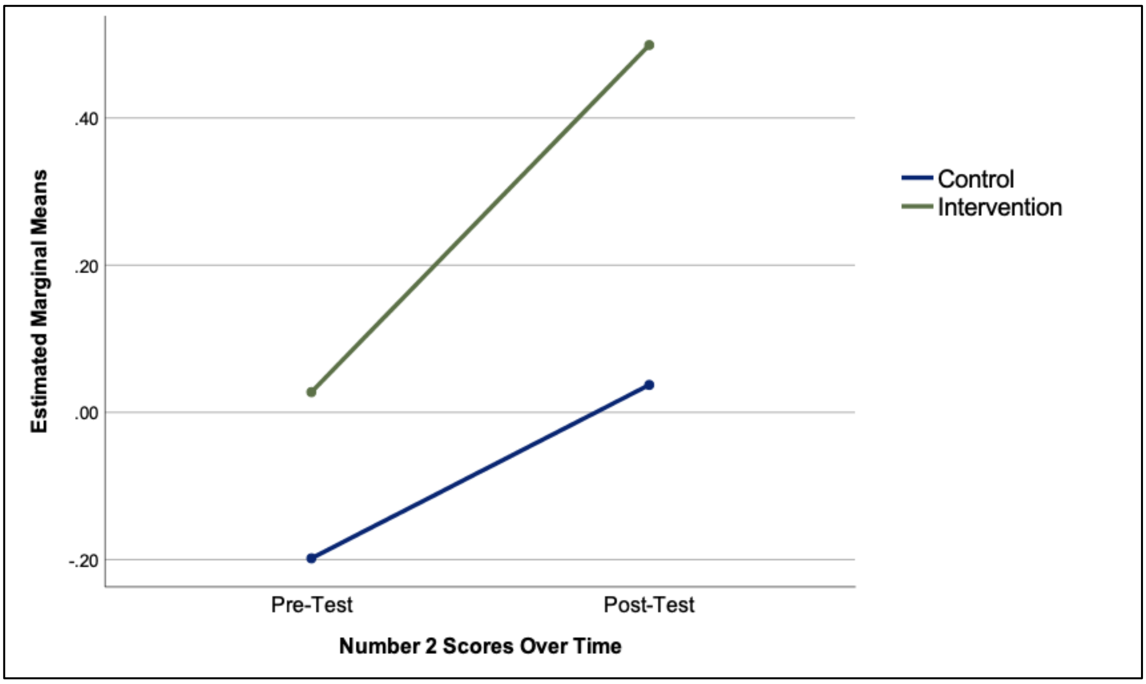


Figure 42: Interactional line chart showing the change in Number 2 scores from the pre-test and post-test assessments in the control (blue line) and intervention (green line) classrooms.

Further analysis of the factors driving the interaction indicates that at the pre-intervention stage there was a significant difference between the two conditions, $t(493)=-2.04$, $p=0.042$. The post-intervention analysis showed a significantly larger difference between the control and intervention classrooms: $t(493)=-3.87$, $p<0.001$. Once again, the mean difference score was higher in the post-test scores (MD=-0.461) compared to the pre-test (MD=-0.226). The between-group effect size in favour of the dynamic soundfield intervention was medium ($\boldsymbol{d}=0.35$). Together, these findings suggest that learners in the dynamic soundfield classrooms improved their skills in formal arithmetic at a higher level than the control condition. However, there was a significant difference between the groups at the pre-trial stage. Interestingly, analysis of the primary 1 PIPS scores (section 6.4) demonstrated there was a significant difference between the groups in numeracy when starting and finishing the first year of school.

6.6.3 Data handling subtest

As Table 46 illustrates, the analysis conducted on the data handling subtest found a significant difference between the two pre-test and post-test assessment time points, which generated a medium effect size. The main effect of condition was also significant indicating that the age equivalent scores were different between the control and intervention classrooms. The magnitude of the difference was small. The condition x time interaction was also significant indicating that any difference in the pre-intervention and post-test scores depended on the classroom the learner attended. The interaction for the data handling subtest is illustrated in Figure 43.

Two paired and two independent *t*-tests (two-tailed) with a Bonferroni correction ($\alpha/4=0.0125$) applied were run to decompose the interaction further. The largest change in scores from baseline to follow-up were found in the intervention classrooms (M=-0.497, SD=1.2) rather than the control (M=-0.235, SD=1.4). Unsurprisingly in a longitudinal study into attainment, both the control, $t(216)=-2.48$, $p=0.014$ and intervention classrooms, $t(277)=-7.29$, $p<0.001$ demonstrated a significant improvement in scores over the academic year. The results suggest that, in comparison to the control group, the learners in the adaptive amplified classrooms gained a beneficial effect of dynamic soundfield on their data handling skills. Furthermore, both the control and intervention classes showed a significant difference at both the pre-intervention, $t(493)=-2.04$, $p=0.042$, and the post-intervention stage, $t(493)=-3.86$, $p<0.001$. The intervention classrooms (M=0.509) once again demonstrated the largest improvement compared to the control (M=0.248). The effect size in favour of the dynamic soundfield classrooms was medium ($\boldsymbol{d}=0.34$). Overall the results suggest a dynamic soundfield advantage in data handling.

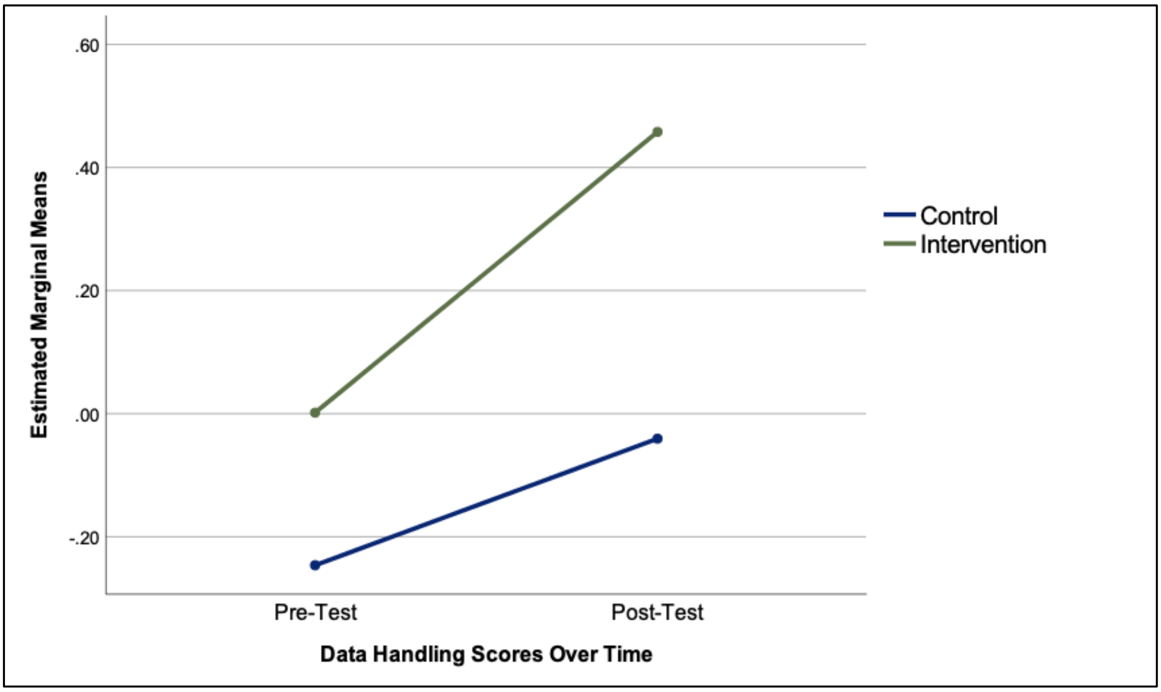


Figure 43: Interactional line chart showing the change in data handling scores from the pre-test and post-test assessments in the control (blue line) and intervention (green line) classrooms.

6.6.4 Measures, Shape and Space subtest

As Table 46 illustrates, the change in Measure, Shape and Space scores was significant throughout the study, but the effect size for this change was small. Paired *t*-tests (two-tailed) revealed there was a significant improvement between assessment time point 1 and time point 2 in both the control, $t(216)=-2.64$, $p=0.009$, and intervention classrooms, $t(277)=-2.87$, $p=0.004$. Inspection of the means shows the change in scores were very similar across both groups: control ($M=-0.227$, $SD=1.3$), intervention ($M=-0.213$, $SD=1.2$). As Table 46 also illustrates there was a significant main effect of condition, which again produced a small effect size. Two independent *t*-tests (two-tailed) showed that there was a significant difference between the groups at the pre-intervention, $t(493)=-3.49$, $p<0.001$, and the post-intervention stage, $t(493)=-2.91$, $p=0.002$. Analysis of the means again indicates the two groups have similar mean differences. Importantly, there was no significant condition x time interaction, and so any changes observed could be attributable to other factors than the intervention.

6.6.5 General Mathematics module

To investigate the change in the learner's standardised General Mathematics scores over time, a two-way mixed ANOVA was conducted. As Table 46 illustrates, there was a significant main effect of time, generating a small effect size. The analysis revealed that both the control, $t(216)=-3.03$, $p=0.003$, and intervention groups, $t(277)=-4.01$, $p<0.001$, improved their standardised scores performance between the pre-test and post-test stages. The mean difference in the control (M=2.19, SD=11) and intervention (M=2.62, SD=11) classrooms were fairly equal, indicating that both groups made similar progress from baseline.

There was also a significant main effect of condition, once again this produced a small effect size. Independent *t*-tests show there was a significant difference between the two groups at the pre-intervention, $t(493)=-4.25$, $p<0.001$, and post-intervention time points, $t(493)=-3.99$, $p<0.001$. The mean difference between the pre-intervention (M=5.33, SD=1.2) and post-intervention (M=5.75, SD=1.4) assessments was fairly similar. There was also no significant condition x time interaction which suggests that the changes observed may be a factor of the sample rather than the intervention. Overall, there is no significant evidence to suggest that the dynamic soundfield system influenced the outcomes in the General Mathematics module.

6.6.6 General Mathematics subtests and C_{50}

Only Number 2 and Data Handling subtests demonstrated a significant time x condition interaction. These tests were repeated controlling for the three different speech weighted C_{50} acoustic conditions. There were no significant interactions for the Number 2 subtest in the excellent, $F(1, 44)=3.69$, $p=0.061$, good, $F(1, 350)=1.62$, $p=0.204$ or fair, $F(1, 95)=0.574$ $p=0.451$ conditions. This suggests that when the different acoustic conditions for speech are considered individually, there was no attributable effect for the dynamic soundfield.

In contrast, the mixed ANOVA on the data handling subtest revealed a significant time x condition interaction in both the excellent, $F(1, 44)=4.69$, $p=0.036$, $\eta_p^2=0.096$ and good C_{50} classrooms, $F(1, 350)=4.66$, $p=0.032$, $\eta_p^2=0.013$. The magnitude of the difference was large in the excellent classrooms and small in the good. As Figure 44 illustrates, in the excellent C_{50} classrooms the group means cross, creating a disordinal interaction. When a disordinal is present any interpretation of the main effects may be misleading and instead, the interactions should only be investigated (Field, 2013).

The driver for the interaction appears to be the significant increase in scores between pre-test and post-test assessment time points in the intervention classrooms (green line) and the modest decrease in scores in the control environment (blue line). Further analysis was performed to decompose the interaction and validate this initial analysis.

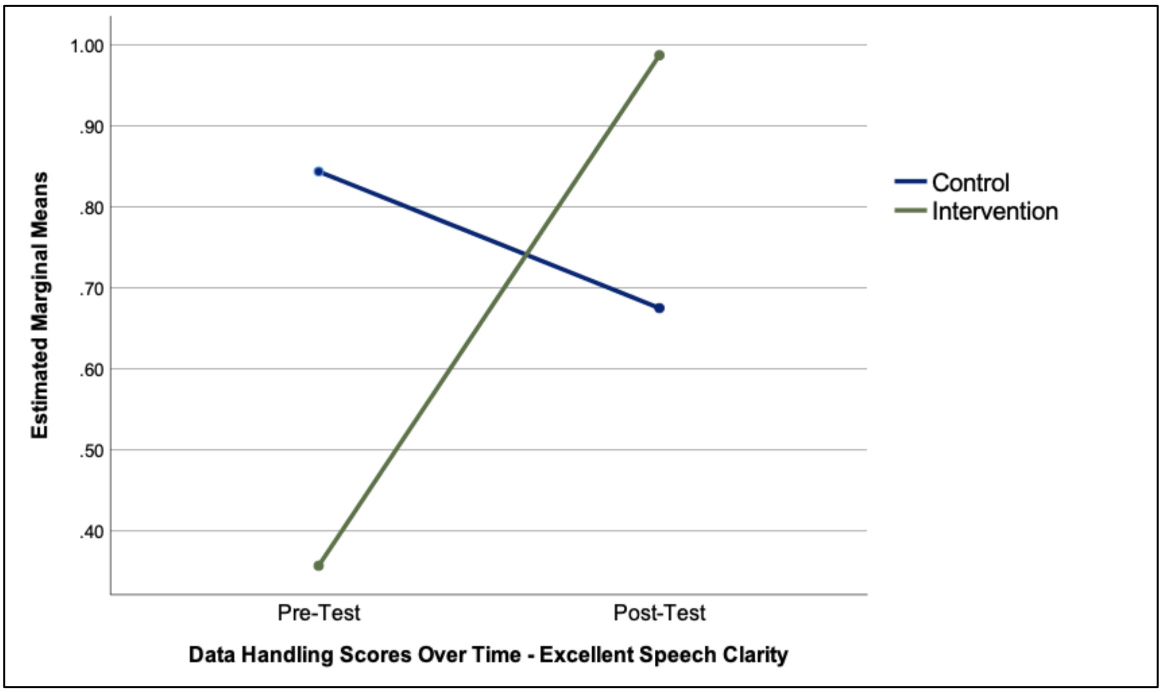


Figure 44: Interactional line chart showing the change in data handling scores from the pre-test and post-test assessments in excellent C_{50} (n=2) classrooms in the control (blue line) and intervention (green line) classrooms

As expected in data handling, the learners in the control classrooms with excellent C_{50} values showed no significant progress over the year of the study, $t(18)=-0.590$, $p=0.562$. In contrast, the classrooms exposed to dynamic soundfield demonstrated significant progress, $t(26)=-2.68$, $p=0.013$. It appears that the change in scores was the primary driver of the interaction. As discussed, this study had only two classrooms categorised as having excellent speech-weighted C_{50} properties, and all the participants were from the least deprived quintile. Therefore, interpretation of the results in terms of drawing conclusions about the effectiveness of dynamic soundfield in excellent C_{50} conditions needs to be treated with caution.

In the good acoustic classrooms for speech, there was a significant main effect of time, $F(1, 350)=27.74$, $p<0.001$, $\eta_p^2=0.073$ which generated a medium effect size. A significant main effect of condition was also identified, $F(1, 350)=20.53$, $p<0.001$, $\eta_p^2=0.055$, which produced a small to moderate effect size. As Figure 45 reveals, there was a significant time x condition interaction, $F(1, 350)=4.66$, $p=0.032$, $\eta_p^2=0.013$. The analysis shows that in the control classrooms, moderate progress was made from baseline to follow-up, but this was non-significant, $t(155)=1.90$, $p=0.06$. In contrast, the intervention classroom showed a significant increase in scores at follow-up, $t(155)=6.07$, $p<0.001$. At the end of the study, as Figure 45 illustrates the gap between the two groups had significantly increased, $t(350)=4.83$, $p<0.001$. Overall, these results suggest that in comparison to the control classrooms, the learners exposed to soundfield in good C_{50} classrooms demonstrated a significant improvement in their knowledge and skills of data handling.

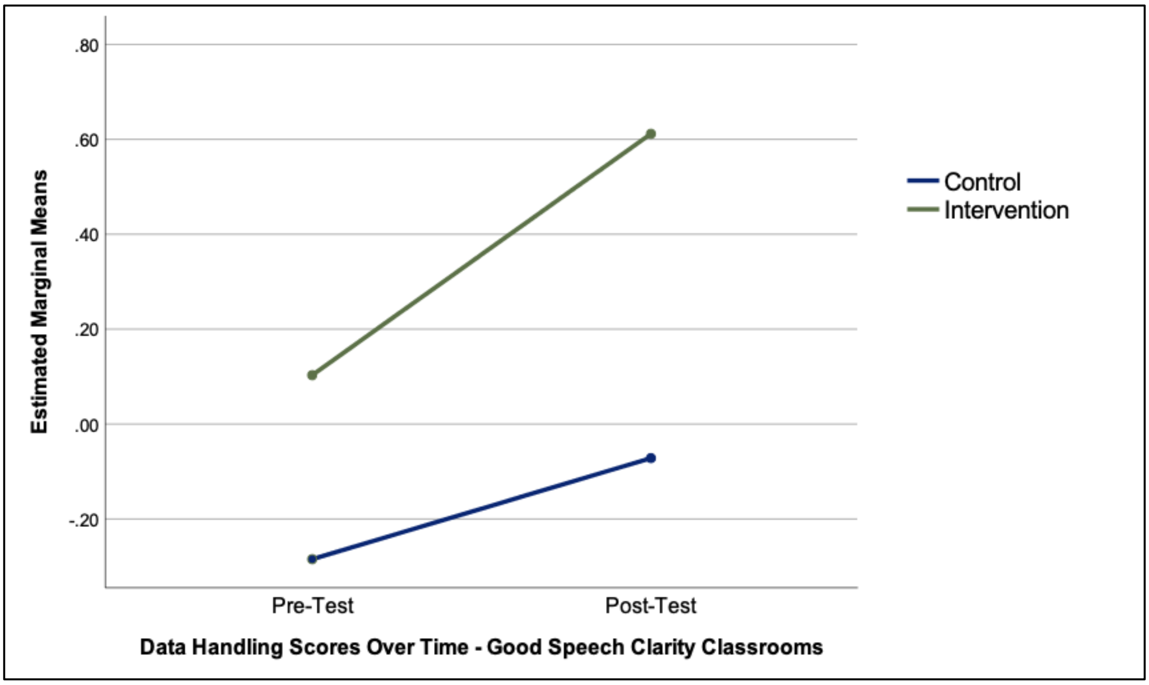


Figure 45: Interactional line chart showing the change in data handling scores from the pre-test and post-test assessments in good C_{50} classrooms (n=18) in the control (blue line) and intervention (green line) classrooms.

6.6.7 Summary – General Mathematics module and subtest

Only subtests Number 2 and Data Handling showed a significant effect of the dynamic soundfield intervention. When accounting for the acoustic conditions, only Data Handling in excellent and good classrooms for speech demonstrated a significant interaction effect. Once again, the results suggest that in comparison to the control, dynamic soundfield is mainly effective in classrooms that have good or excellent acoustics for speech. However, due to the small sample size caution needs to be applied to the results in the excellent C_{50} classrooms.

6.7 AfE Assessments - Mental Arithmetic module

To compare the changes over time, a two-way mixed ANOVA was conducted on the standardised scores for the Mental Arithmetic module. The analysis revealed a significant main effect of time, $F(1, 493)=35.16$, $p<0.001$ $\eta_p^2=0.067$ with a medium effect size. There was also a significant main effect of condition, $F(1, 493)=8.93$, $p=0.003$ $\eta_p^2=0.018$. This suggests that if time (pre and post-test) was not accounted for then there were significant differences between the control and intervention groups. The effect size for the difference between the two groups was small. There was also a significant interaction between condition and time, $F(1, 493)=3.98$, $p=0.047$ $\eta_p^2=0.008$, indicating that changes in the mental arithmetic standardised scores in the group exposed to dynamic soundfield were significantly different to the control classrooms (see Figure 46).

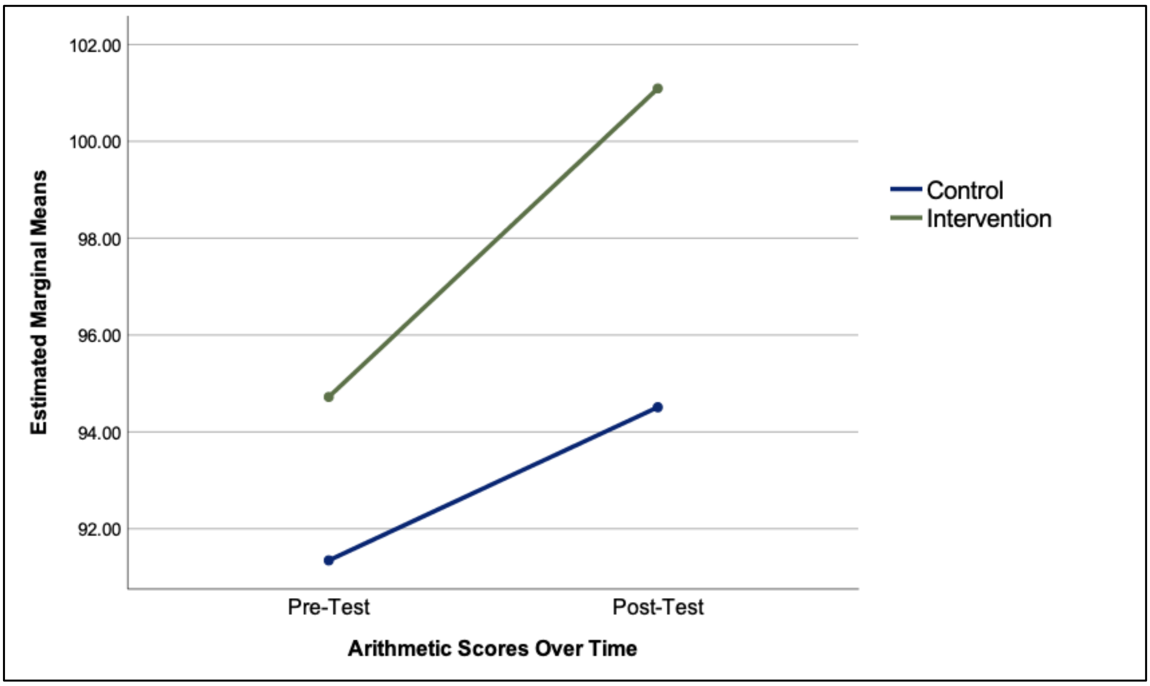


Figure 46: Interactional line chart showing the change in mental arithmetic scores from the pre-test and post-test assessments in the control (blue line) and intervention (green line) classrooms.

Paired *t*-tests (two-tailed), revealed that both the control, $t(216)=-2.55$, $p=0.011$ and intervention classrooms, $t(277)=-6.13$, $p<0.001$ significantly improved mental arithmetic scores between the pre-test and follow-up stages. Analysis of the mean differences show that the largest improvement was in the intervention classrooms (MD=-6.37, SD=17.3) exposed to the dynamic soundfield system compared to the control (MD=3.16, SD=18.2) An independent *t*-test (two-tailed) on the mental arithmetic standardised scores also showed a non-significant difference at the pre-trial stage, $t(493)=-1.93$, $p=0.06$. At the end of the intervention there was a significant difference between the classrooms exposed to dynamic soundfield and the control, $t(493)=-3.38$, $p<0.001$. Analysis of mean difference values shows that the post-test scores were higher. The effect size in favour of the intervention classrooms was medium ($d=0.31$). The current study extends our understanding in this area as the results suggest that amplification of the teacher's voice appears to provide a dynamic soundfield advantage in mental arithmetic, a task that relies heavily on working memory (Cragg et al., 2017).

6.7.1 Mental Arithmetic module and C_{50}

Only classrooms categorised as having good C_{50} measurements demonstrated a positive effect for the dynamic soundfield amplification. There was a significant main effect of time between the pre-test and post-test scores, $F(1, 350)=38.44$, $p<0.001$, $\eta_p^2=0.099$, this generated a large effect size. There was also a significant main effect of condition, $F(1, 350)=11.80$, $p<0.001$, $\eta_p^2=0.033$. The magnitude of the difference between the groups was small. The interaction between time x condition in the good C_{50} acoustic environments also achieved significance, $F(1, 350)=6.77$, $p=0.010$, $\eta_p^2=0.019$. As Figure 47 shows there was a significant difference in the outcomes observed between the two groups in mental arithmetic. The intervention classrooms (green line) demonstrated a significant improvement in scores from baseline. In contrast to the control group (blue line) recorded more modest progress. Interestingly, the configuration of Figure 47 is fairly similar to Figure 46 suggesting that the improved scores observed in section 6.7 were primarily the result of a good acoustic environment.

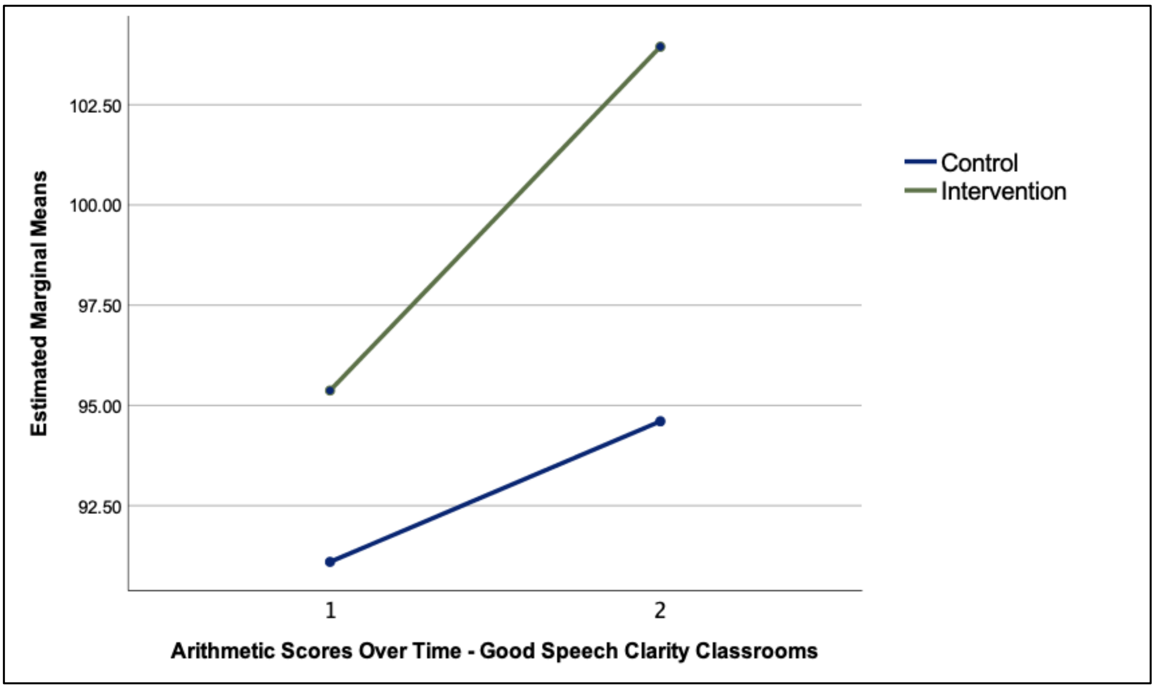


Figure 47: Interactional line chart showing the change in mental arithmetic scores from the pre-test and post-test assessments in good speech-weighted C_{50} classrooms (n=18) in the control (blue line) and intervention (green line) classrooms.

Further statistical analysis was carried out to explore the interaction. Both the control, *t*(155)=-2.33, p=0.021 and intervention, *t*(195)=-6.81, p<0.001 classrooms demonstrated significant improvement over the time of the study. As anticipated, the change in mean scores was higher in the intervention classrooms (M=-8.58, SD=18) compared to the control (M=-3.51, SD=18). Independent *t*-tests showed at the pre-intervention stage there was a significant difference between the groups, *t*(350)=-2.02, p=0.044. At the end of the study, there was also a significant difference, *t*(350)=-4.06, p=<0.001. The *p*-values indicate that there was a more significant effect at the post-test stage compared to baseline. Unsurprisingly, an inspection of the mean differences confirms that there was a larger change in scores at the post-intervention stage (M=9.34) compared to baseline (M=4.3). The between-group effect size at the end of the intervention in favour of the soundfield classrooms was medium ($\boldsymbol{d}$=0.43). Overall the results suggest that classrooms with good C_{50}

measures achieve a significant benefit from dynamic soundfield amplification in mental arithmetic.

6.8 AfE assessments – Reading and Spelling

Data were analysed using a two-way mixed ANOVA with time and condition as the independent variables and reading standardised scores and age differences subtest scores as the dependent variable. Table 47 presents the results.

	Within-Subject			Between-Subject			Interactions		
α=0.05 (1,493)	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
Word Recognition	127.54	0.001	0.206	13.11	0.001	0.026	23.47	0.001	0.045
Word Decoding	121.95	0.001	0.198	11.82	0.001	0.023	10.67	0.001	0.021
Comp	123.66	0.001	0.201	14.67	0.001	0.029	20.02	0.001	0.039
Spelling	209.35	0.001	0.298	8.47	0.004	0.17	1.69	0.194	0.003
Reading	180.43	0.001	0.296	14.73	0.001	0.029	29.80	0.001	0.065

Table 47: Results from the two-way mixed ANOVAs conducted on the Reading module and subtests.

6.8.1 Word Recognition subtest

The descriptive data for the Word Recognition subtest shows that at baseline the control group were approximately seven months behind their age expected level and the intervention classes were approximately four months behind. Importantly, at follow-up, the progress recorded was of a greater magnitude in the classrooms exposed to dynamic amplification (approximately nine months) compared to the control (approximately four months). As Table 47 illustrates, the improved scores from baseline are reflected in a significant main effect of time which generated a large effect size. Unsurprisingly, there was also a significant main effect of the between-subjects factor, condition. As Figure 48 illustrates, the condition x time interaction was also significant indicating that the improvement in word recognition scores in the classrooms exposed to dynamic soundfield was significantly different than the control.

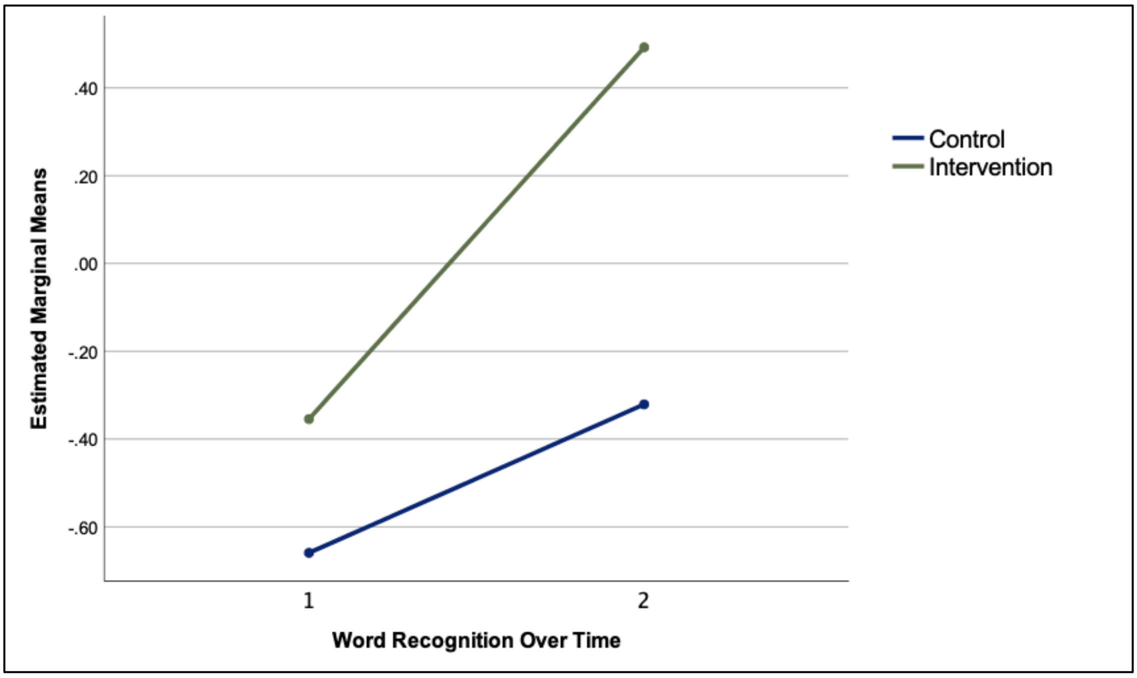


Figure 48: Interactional line chart showing the change in word recognition scores from the pre-test and post-test assessments in the control (blue line) and intervention (green line) classrooms.

Paired *t*-tests (two-tailed), confirmed that in both groups, learner's performance improved from the start of the academic year to the end: control, $t(216)=-4.16$, $p<0.001$ and intervention, $t(277)=-14.23$, $p<0.001$. The mean difference scores were higher in the intervention classroom (M=-0.847, SD=1) compared to the control (M=-0.338, SD=1.4). The results indicate that although there was an improvement over time for both conditions, as would be expected in a longitudinal study on academic achievement there was a significant differential effect for learners in the dynamic soundfield classes. Consistent with these results, independent *t*-tests (two-tailed) identified that there was a non-significant difference between the groups at the pre-intervention stage, $t(493)=-1.93$, $p=0.06$. As Figure 48 demonstrates the change in scores resulted in a significant difference between the groups at follow-up, $t(493)=-4.77$, $p<0.001$. The effect size in favour of the intervention classrooms over the control was medium ($\boldsymbol{d}=0.43$). Overall, the results suggest that the intervention classrooms gained a significant beneficial effect from adaptive amplification.

6.8.2 Word Decoding subtest

Consistent with the results from the Word Recognition subtest, the main effect of time was significant in the Word Decoding subtest. Across both the control and intervention classrooms word decoding skills improved at the end of the primary 3 year compared to the start of term. The effect size for the change in scores was large. As Table 47 shows the main effect of condition was also significant, indicating there was a significant difference in word decoding skills between the control and intervention classrooms. The effect size for the difference was small. The time x condition interaction was also significant, indicating that the level of one factor depends on the level of the other factor. Once again, the effect size for the interaction was small. As Figure 49 illustrates the interaction appears to be primarily driven by the differential change in post-test scores in the dynamic soundfield classrooms.

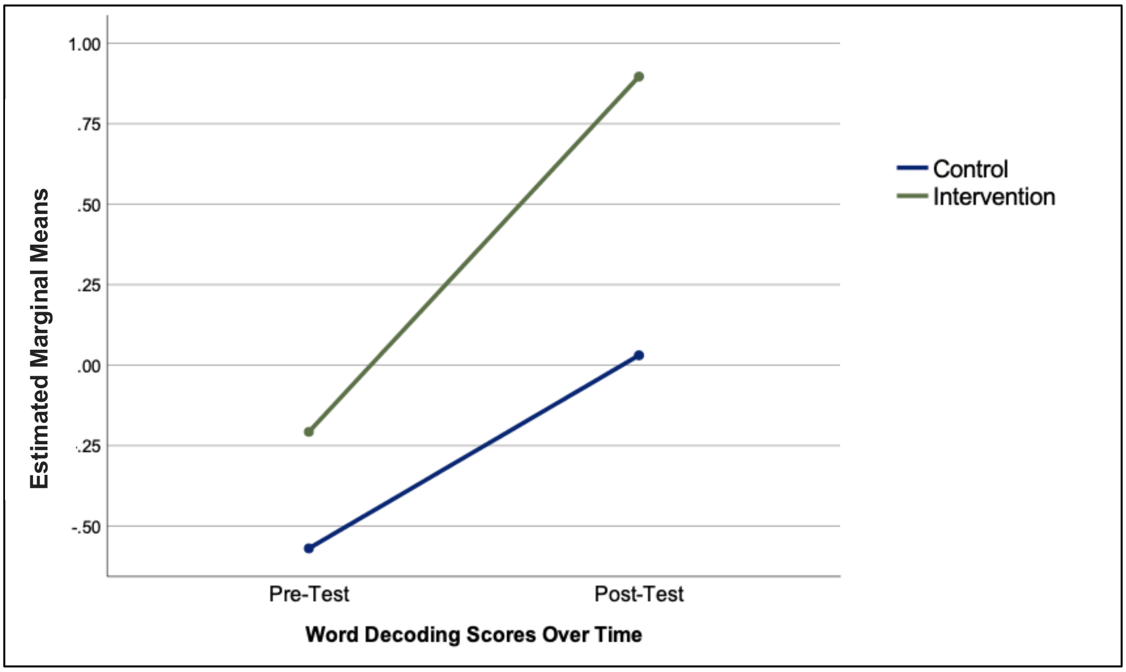


Figure 49: Interactional line chart showing the change in word decoding scores from the pre-test and post-test assessments in the control (blue line) and intervention (green line) classrooms.

In order to confirm the factors driving the interaction, further analysis was undertaken. Paired *t*-tests (two-tailed), revealed that in both groups there was a significant increase in word decoding scores from the pre-test and post-test assessments: control, $t(216)=-5.44$, $p<0.001$ and intervention, $t(277)=-10.44$, $p<0.001$. Inspection of the means reveals that the change in scores was higher in the intervention classrooms (M=-1.10, SD=1.7) compared to the control (M=-0.599, SD=1.6). In addition, the independent *t*-tests (two-tailed) showed a non-significant difference between the groups at baseline, $t(493)=-1.93$, $p=0.06$. At follow-up there was a significant difference between the groups, $t(493)=-4.30$, $p<0.001$. The magnitude of the difference was medium ($d=0.44$). Overall, the results suggest that the change in scores in the dynamic soundfield classrooms drove the interaction.

6.8.3 Reading Comprehension subtest

As discussed previously, learners only access the reading comprehension model if their combined word decoding and word recognition score add up to 8.0 or more. Due to the methods used to calculate the overall reading score the comprehension age equivalent score has a floor of 4.0 to compensate for missing data. Therefore, interpretation of the results for the reading comprehension subtest in terms of drawing conclusions about the effectiveness of the dynamic soundfield intervention has to be treated with caution.

As Table 47 shows, there was a significant main effect of time. The magnitude of the effect size between assessment time points 1 and 2 was large. There was also a significant main effect of condition. The effect size between the two groups was small. The condition x time interaction was also significant, this suggests that there was a significant difference in how the control and intervention classrooms responded over the time of the study. Figure 50 provides a graphical representation of the interaction showing the change in scores between time points 1 and 2 in both the control and intervention classrooms.

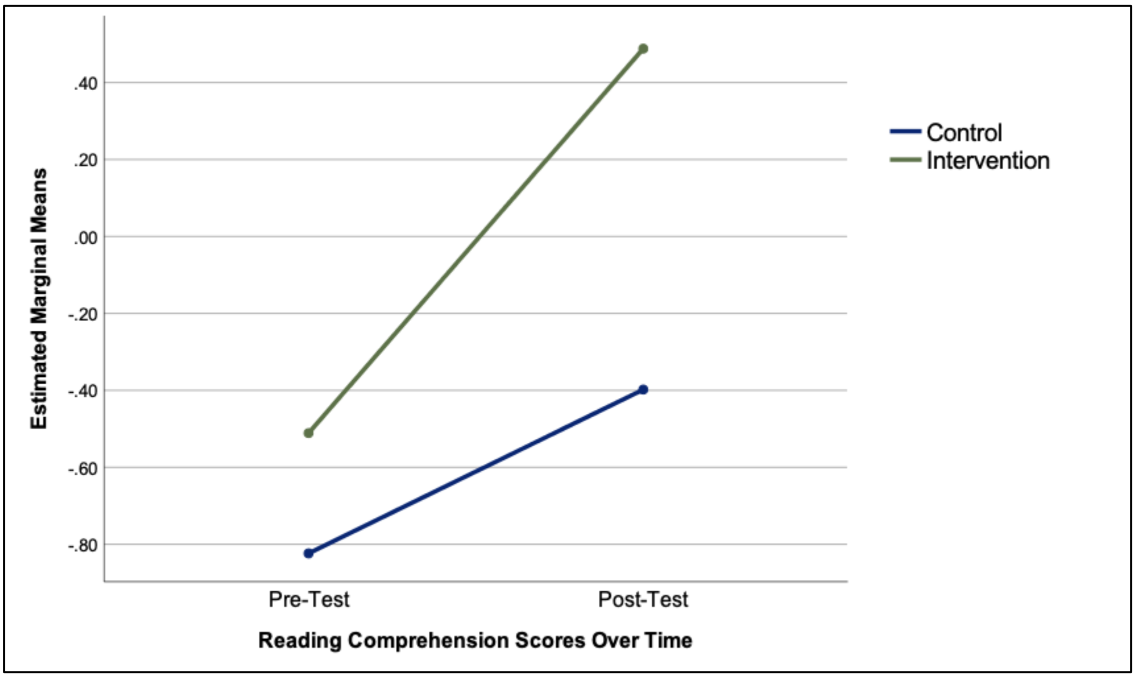


Figure 50: Interaction line chart showing the change in reading comprehension scores from the pre-test and post-test assessments in the control (blue line) and intervention (green line) classrooms.

The significance of the interaction was explored using four *t*-tests with a Bonferroni correction applied ($\alpha/4 = 0.0125$). The Paired *t*-tests (two-tailed), revealed that the learners in both the control and intervention classrooms significantly improved their scores in reading comprehension between the pre-test and post-test stages: control, $t(216)=-4.23$, $p<0.001$ and intervention, $t(277)=-12.25$, $p<0.001$. The mean increase in scores was higher in the intervention classrooms (M=-0.998, SD=1.3) compared to the control (M=-0.425, SD=1.5). This suggests that in comparison to the control, the intervention classrooms made significantly greater improvements in reading comprehension. The independent *t*-tests (two-tailed) showed a non-significant difference between the control and intervention classrooms before the installation of the dynamic soundfield system, $t(493)=-1.91$, $p=0.06$. By the end of the study, there was a significant difference between the control and intervention groups), $t(493)=-5.09$, $p<0.001$. This suggests that learners exposed to dynamic soundfield amplification showed significant improvement in reading comprehension skills compared to the control group. The effect size was small to medium ($\boldsymbol{d}=0.46$).

6.8.4 Spelling subtest

Analysis of the Spelling results showed there was a significant main effect of time. The effect size for the difference between assessment time point 1 and time point 2 was large. Both the control, $t(216)=-9.26$, $p<0.001$ and intervention, $t(277)=-11.47$, $p<0.001$ classrooms demonstrated significant progress during the academic year. The mean increase in scores was modestly higher in the intervention classrooms ($M=-0.818$, $SD=1.2$) compared to the control ($M=-0.683$, $SD=1.1$). There was a significant main effect for condition. The analysis revealed that there was a significant difference between the groups at baseline, $t(493)=-2.48$, $p=0.014$. At the end of the study, the results were also significantly different, $t(493)=-2.98$, $p=0.003$. The effect size was small. There was no significant interaction between condition and time and so the changes observed cannot be attributable to the intervention. No further testing was performed.

6.8.5 Reading module

There was a significant main effect of time, which generated a large effect size. As Table 47 illustrates there was also a significant main effect of condition, with a small effect size. Of most importance, there was a significant interaction between condition and time. The magnitude of the effect size for the interaction was medium. (see Figure 51).

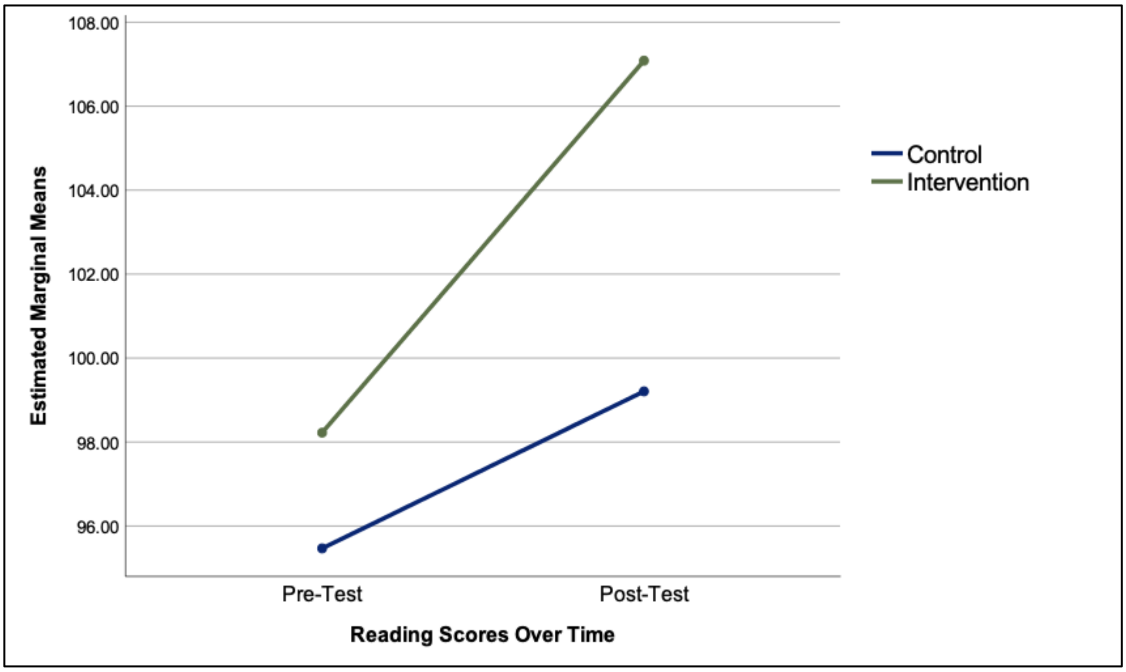


Figure 51: Interaction line chart showing the change in reading scores from the pre-test and post-test assessments in the control (blue line) and intervention (green line) classrooms.

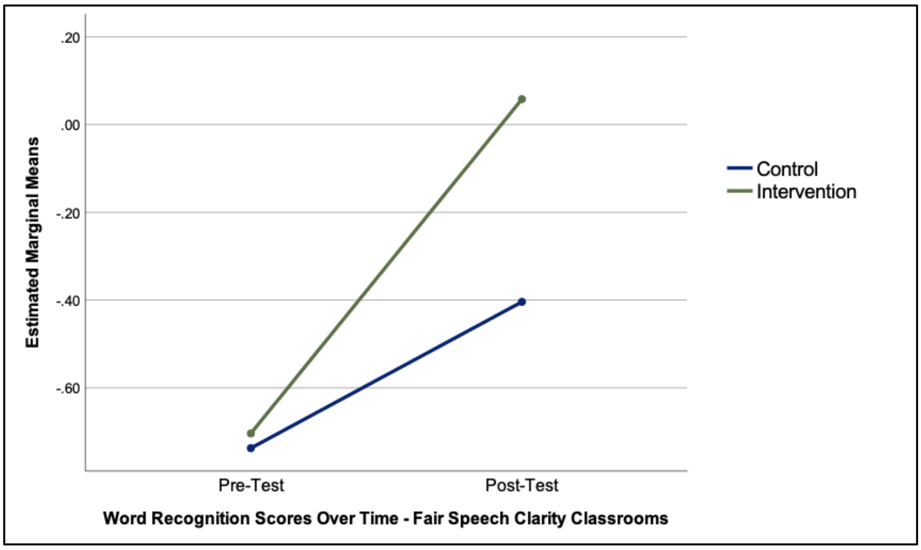
Paired *t*-tests (two-tailed), revealed that both the control classrooms, $t(216)=-5.25$, $p<0.001$ and intervention classrooms, $t(277)=-16.50$, $p<0.001$ significantly improved their standardised scores between the pre-test and post-test assessments. Independent *t*-tests (two-tailed) also showed a non-significant difference between the groups at baseline, $t(493)=-1.96$, $p=0.06$. In contrast, at the follow-up stage there was a significant difference between the control and intervention classrooms, $t(493)=-5.20$, $p=<0.001$. Analysis of the mean differences shows that the largest improvement was in the intervention classrooms (MD=-8.75) exposed to the dynamic soundfield system compared to the control (MD=-3.74). This suggests that the learners exposed to dynamic soundfield increased their overall reading scores at a higher rate than the control group. The effect size in favour of the intervention classrooms was medium ($\boldsymbol{d}=0.46$). This suggests that these results, collectively, produced the interaction effect.

6.8.6 Reading and Spelling module and subtests and C_{50}

Statistical analysis was undertaken to explore any relationship between the speech-weighted C_{50} properties of the classrooms and the improvements observed in the reading module and subtests in the intervention classrooms. The two-way mixed ANOVA was repeated controlling for the three different speech weighted C_{50} classroom categorisations.

The Word Recognition subtest showed a positive time by condition interaction in both the fair, $F(1,95)=6.96$, $p<0.001$, $\eta_p^2=0.068$ and good, $F(1,350)=16.18$ $p<0.001$, $\eta_p^2=0.044$ classroom environments. The magnitude of the effect size was medium in the fair C_{50} classrooms and small in the good. Further analysis was run to decompose the interaction to identify the factors driving it. The interactions illustrated in Figure 52 reveals a significant main effect of time in both the fair, $F(1,95)=45.49$, $p<.001$, $\eta_p^2=0.324$ and good, $F(1,350)=78.04$ $p<0.001$, $\eta_p^2=0.182$ speech-weighted C_{50} classrooms. This demonstrates that the follow-up assessment scores were significantly higher in the intervention classrooms (green line) compared to the control (blue line). Paired *t*-tests confirm that in the fair C_{50} classrooms, both the control, $t(41)=-2.46$, $p=0.018$ and intervention, $t(54)=-7.85$, $p<0.001$ groups made significant progress over the year. The intervention classrooms demonstrated a higher level of change ($M=-0.761$, $SD=0.7$) compared to the control ($M=-0.333$, $SD=0.9$). This pattern was mirrored in the good C_{50} classrooms: control, $t(155)=-2.77$, $p=0.006$ and intervention, $t(195)=-11.41$, $p<0.001$. Once again, the mean difference scores were higher in the intervention ($M=-0.870$, $SD=1.1$) classrooms compared to the control ($M=-0.326$, $SD=1.5$).

A)



B)

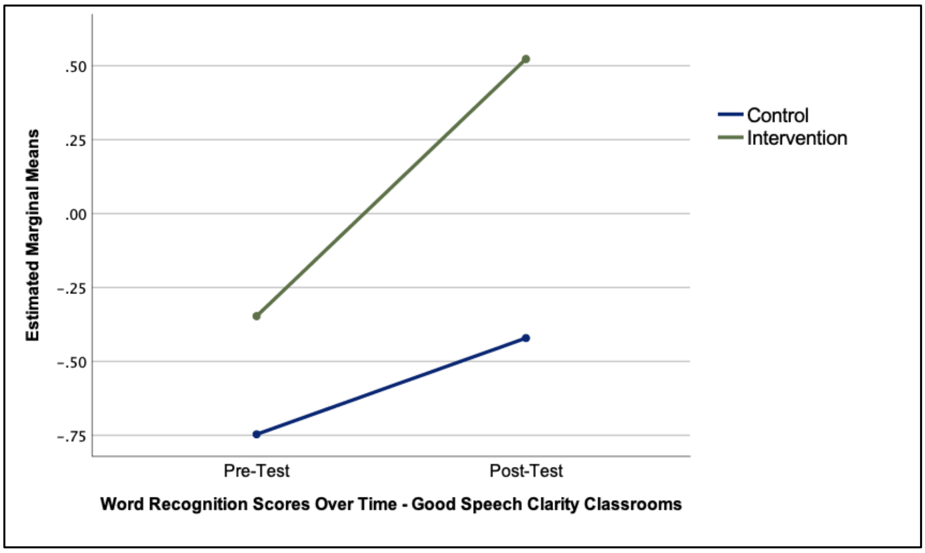


Figure 52: A) Time x condition interaction for the Word Recognition subtest in Fair C_{50} classrooms (n=5). B) ANOVA repeated in good C_{50} classrooms (n=18).

Further analysis showed that in the fair classrooms there was a non-significant difference between the groups at the pre-intervention, $t(95)=-0.107$, $p=0.915$ and follow-up, $t(95)=-1.32$, $p=0.191$ assessment points. In the good C_{50} classrooms, a significant between-subject factor was observed at baseline, $t(350)=-2.11$, $p=0.035$

and follow-up, $t(350)=-4.67$, $p<0.001$. Overall, this suggests that the change in scores was the primary factor driving the interaction.

In the Word Decoding subtest, there was only a positive dynamic soundfield effect in the good classrooms for speech. The mixed ANOVA revealed a significant main effect of time $F(1,350)=82.67$ $p<0.001$, $\eta_p^2=0.191$, this generated a large effect size. A significant main effect of condition was also observed, $F(1,350)=13.10$ $p<0.001$, $\eta_p^2=0.036$. A significant interaction between the pre/post-test scores over time in the classrooms that the learners attended was also revealed $F(1,350)=8.83$ $p=0.003$, $\eta_p^2=0.025$. This interaction is evident in Figure 53, where the change in scores over time is shown in both groups. The intervention classrooms (green line) showed more progress from baseline compared to the control (blue line).

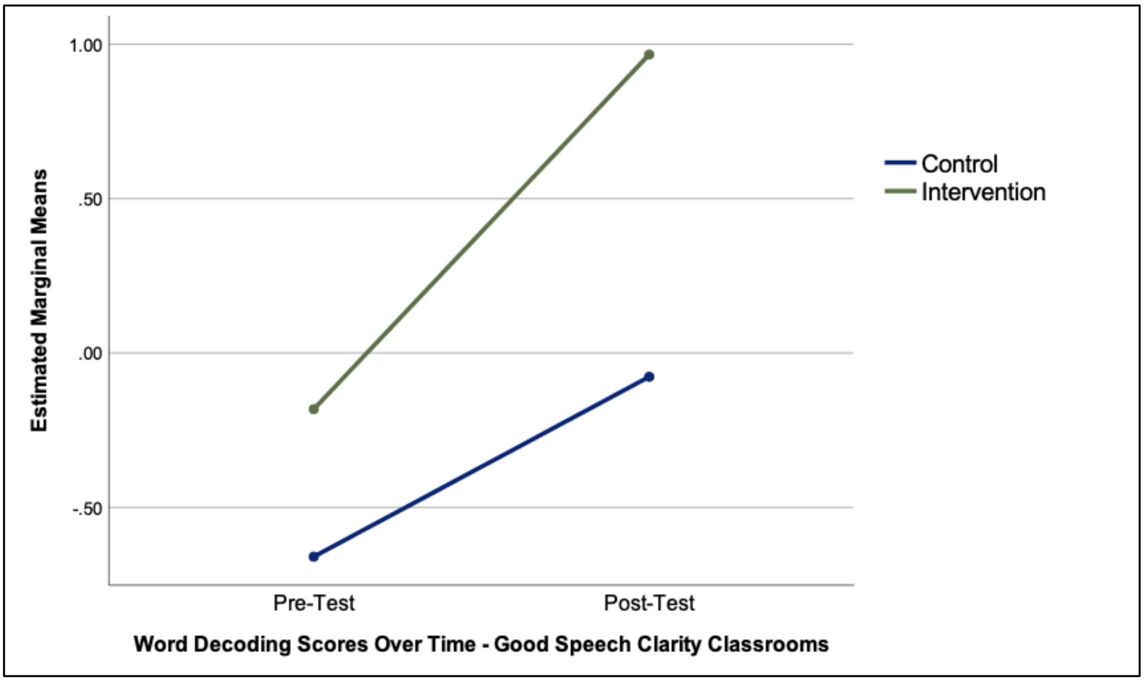


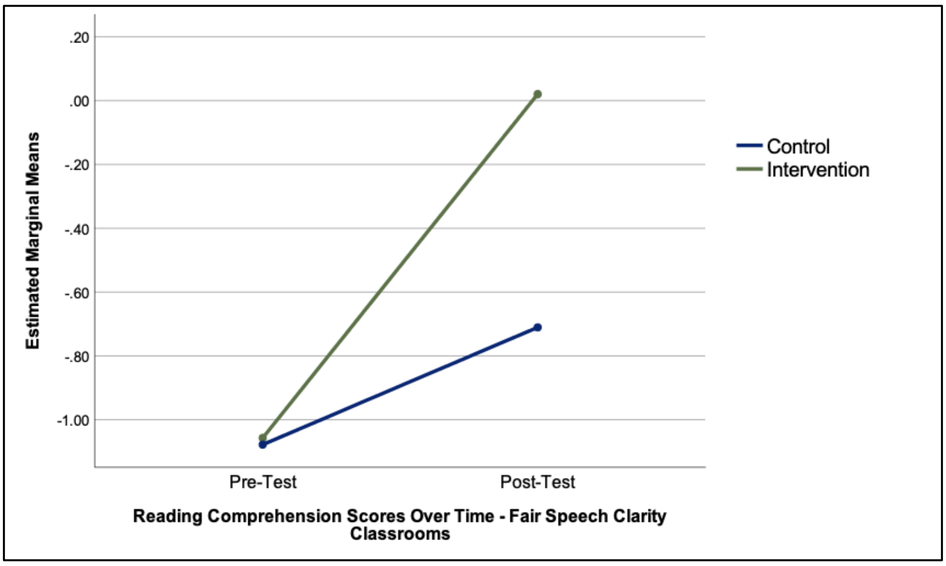
Figure 53: Time x condition interaction for the Word Decoding subtest in Good C_{50} classrooms (n=18).

Analysis was undertaken to discern the independent variables driving the interaction. Paired *t*-tests (two-tailed), revealed there was significant progress during the year in the control, $t(155)=-4.37$, $p<0.001$ and intervention classrooms, $t(195)=-8.66$, $p<0.001$. Inspection of the means confirms that the intervention group ($M=1.15$, $SD=1.8$) made the greatest progress, in comparison to the control, ($M=0.583$, $SD=1.7$). With the Bonferroni correction applied ($\alpha/4=0.0125$), in the good C_{50} classrooms there was a non-significant difference at baseline, $t(350)=-2.18$, $p=0.030$. At follow-up there was a significant difference between the groups, $t(350)=-4.32$, $p<0.001$. Inspection of the *p* values and the mean difference score indicates that the biggest change was at the post-intervention stage ($M=-1.04$) compared to the baseline ($M=-0.478$). The results suggest that the highly significant increase in test scores in the good C_{50} classrooms exposed to dynamic soundfield drove the interaction.

Reading Comprehension in the fair acoustic classrooms, $F(1,95)=6.24$ $p=0.014$, $\eta_p^2=0.062$ and the good acoustic classroom environments for speech clarity, $F(1,350)=11.37$ $p<0.001$, $\eta_p^2=0.031$ demonstrated a significant interaction. As Figure 54 shows the interaction appears to be driven by the change of scores in the intervention classrooms increasing over time more than the control. The analysis confirmed that in the fair classrooms the control group made non-significant progress from baseline, $t(41)=-1.59$, $p=0.12$, in contrast, the classrooms exposed to dynamic soundfield demonstrated a significant improvement in their tariff score, $t(54)=-6.18$, $p<0.001$. In the good speech weighted C_{50} classrooms, both the control, $t(155)=-3.50$, $p<0.001$ and intervention classrooms, $t(195)=-9.63$, $p<0.001$ showed progress over time. Inspection of the means indicates that the progress was at a higher level in the intervention classrooms ($M=-0.951$, $SD=1.4$) compared to the control ($M=-0.427$, $SD=1.5$).

Applying the Bonferroni correction ($\alpha/4=0.0125$), further analysis confirmed that in classrooms with good acoustics for speech there was a non-significant difference between the groups at baseline, $t(350)=-2.23$, $p=0.027$. At follow-up there was a significant difference between the control and intervention groups, $t(350)=-4.62$, $p<0.001$. As anticipated the mean difference was higher in the post-test scores (M=-0.958) compared to the pre-test scores (-0.435), suggesting that the change in post-test scores between the groups contributed to the interaction. In the fair classrooms for speech, there was a non-significant difference between the groups at baseline, $t(95)=-0.064$, $p=0.949$ which although reduced considerably over time remained non-significant, $t(95)=--1.89$, $p=0.062$. Inspection of the mean differences shows that at the pre-test assessment there was a marginal difference between the groups (M=-0.021). At post-test, the difference had increased (M=-0.731). This again suggests that the improved performance in the adaptive amplified classrooms produced a significant interaction effect.

A)



B)

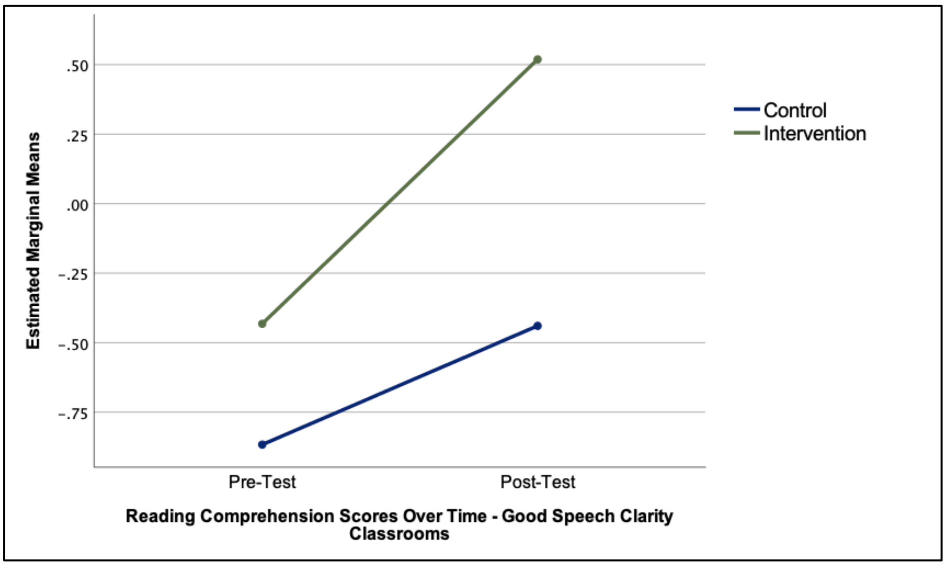
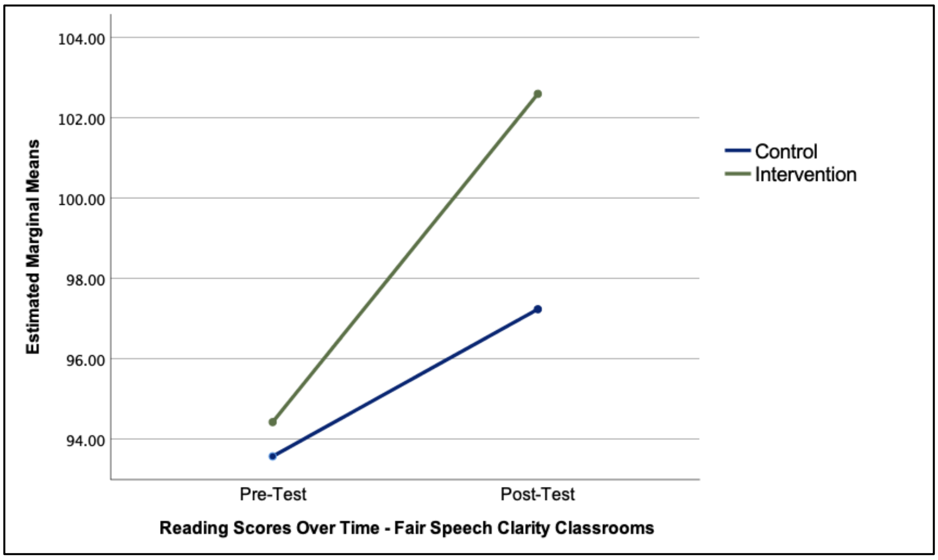


Figure 54: A) Time x condition interaction for the Reading Comprehension subtest in Fair C_{50} (n=5) classrooms. B) ANOVA repeated in good speech clarity classrooms (n=18).

Changes in the pre-test/post-test standardised reading scores over time were found to be highly significant in all three acoustic environments: excellent, $F(1,44)=32.16$ $p<0.001$, $\eta_p^2=0.422$; good, $F(1,350)=132.43$ $p<0.001$, $\eta_p^2=0.274$ and fair $F(1,95)=47.47$ $p<0.001$, $\eta_p^2=0.333$. No significant interaction was observed in the excellent classrooms, $F(1,44)=3.91$ $p=0.06$. As discussed previously, due to the small and unrepresentative sample of learners that attended the excellent classrooms, interpretation of the results in terms of the effectiveness of dynamic soundfield in such conditions should be treated with caution. As anticipated, based on the subtest results a significant interaction with a medium effect size was observed in both the fair, $F(1,95)=6.89$ $p=0.010$, $\eta_p^2=0.66$ and, good classrooms, $F(1,350)=23.44$ $p<0.001$, $\eta_p^2=0.066$. This is illustrated in Figure 55, with the intervention classrooms (green line) producing a steeper trajectory in comparison to the control (blue lines).

A)



B)

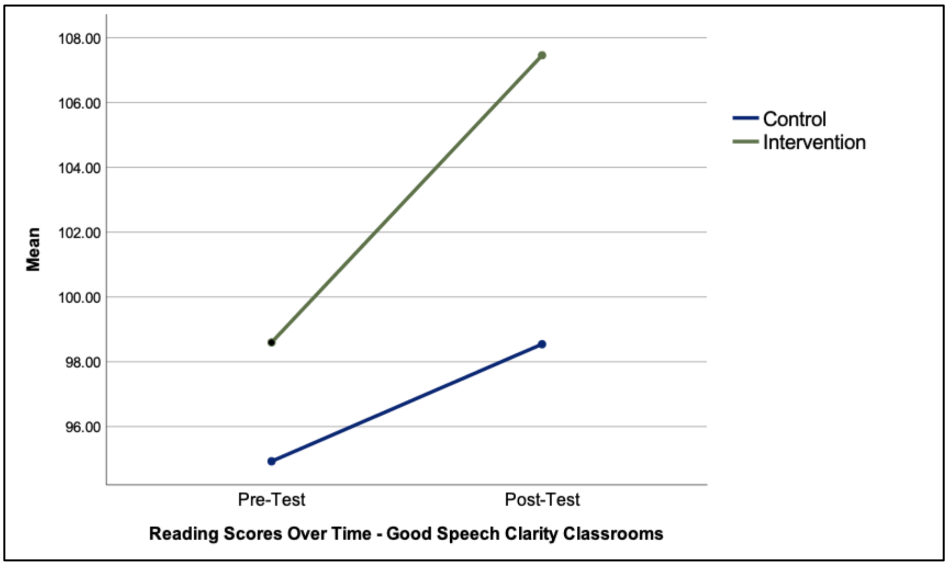


Figure 55: A) Time x condition interaction for the Reading module in Fair C_{50} classrooms (n=5). B) ANOVA repeated in good speech clarity classrooms (n=18).

To investigate the interaction further a series of paired and independent *t*-tests were performed. In the fair classrooms both the control group, *t*(41)=-2.57, p=0.014 and intervention group, *t*(54)=-7.90, p<0.001 demonstrated progress over the time of the study. Inspection of the means confirmed that the intervention classrooms (M=-8.17, SD=7.7) made more progress than the control (M=-3.67, SD=9.23). This appears to be the primary factor in driving the interaction. In the good classrooms for speech, once again both the control, $t(155)=-4.09$, $p<0.001$ and intervention classrooms $t(195)=-13.35$, $p<0.001$ made progress. The mean scores reveal that not only did the intervention classes (M=-8.86, SD=9.3) outperform the control, (M=-3.61, SD=11.03) but that the rate of progress was similar to the fair classroom environments. With the Bonferroni correction applied ($\alpha/4 = 0.0125$), the difference between the groups was non-significant at the start, $t(350)=-2.19$, $p=0.028$ and significant at the end, $t(350)=-4.92$, $p<0.001$ of the study. Once again, the differential change in the pre-test/post-test scores between the control and intervention groups generated the interaction.

6.8.7 Summary – Reading module and subtest

The three main components of reading - word recognition, word decoding, and reading comprehension - demonstrated a significant interaction effect. Follow-up analysis confirmed that this was primarily driven by the scores in the classrooms exposed to dynamic soundfield increasing at a higher rate compared to the control. Interestingly, when comparing the different acoustic conditions, the intervention classrooms categorised as fair or good for speech clarity demonstrated significantly higher scores in word recognition, reading comprehension and reading compared to the control. Only the good C_{50} intervention classrooms demonstrated a significant improvement in word decoding.

6.9 Discussion of AfE results

The findings relating to improved performance in curricular areas associated with both verbal and non-verbal processing support previous studies which suggest that young people require an enhanced SNR to ensure appropriate levels of speech intelligibility (Elliott, 1979, Boothroyd, 2004, Bradley and Sato, 2004). One of the potential benefits of the dynamic soundfield system is that it provides a variable SNR by monitoring the level of background noise and adapting the amount of gain provided to the teacher's voice. The dynamic adaptation is triggered when background noise levels reach 54dB SPL (Phonak, 2014). Dance et al. (2018) measured the effectiveness of the dynamic soundfield system under different room conditions (varying ambient noise pressure and reverberation times) using the equivalent noise reduction (ENR) criterion. The ENR is in effect the inverse of SNR, where the improvements in speech intelligibility by the dynamic soundfield are equated to what would be achieved by a reduction in the levels of background noise. The results suggest that when interfering noise levels are 40dBA the dynamic soundfield system provided no discernible benefit to speech intelligibility. As the level of competing noise increases, so does the effectiveness of the system at improving speech intelligibility. Saturation is reached at 7.7dB(A) ENR with background noise levels above 75dB(A).

The results in Chapter 5 demonstrated that the mean noise levels in the three curricular areas observed was 64.2dB(A) L_{Aeq} in the control and 64.53dB(A) L_{Aeq} in the intervention classrooms. The assessment of the acoustical quality of the research classrooms was measured in Chapter 4 with a $T_{mf\ mean}$ of 0.76s (SD=0.17s) recorded in the intervention classrooms. In the experimental study by Dance et al. (2018) rooms with an RT_{60}=0.8s fitted with a dynamic soundfield increased speech intelligibility when noise levels were ≥44dB(A). The estimated equivalent noise reduction was >6dB. For classrooms with longer reverberation times (RT≥1.0s) the dynamic soundfield improved speech intelligibility when noise levels were ≥50dB(A) and for shorter reverberation times (RT≤0.6s) the noise levels were ≥38dB(A). In Chapter 5 the lowest noise levels recorded in all curricular areas were above the minimum noise threshold levels that dynamic soundfield becomes effective: numeracy 58.9 L_{Aeq}, literacy 57.2 L_{Aeq} and IDL 57.6 L_{Aeq}. Taken together, it would appear that dynamic soundfield is effective for improving speech intelligibility in a range of reverberant times at noise levels commonly observed in Scottish primary 3 classrooms. This is a new finding, not revealed in any study before.

Comparisons on the effectiveness of dynamic soundfield based on the three different acoustic classroom categorisations suggest that dynamic soundfield is primarily effective in classrooms with good or excellent C_{50} values in subjects associated with non-verbal tasks (numeracy and developed ability). This is consistent with previous studies that have shown the importance of early reflections (C_{50}) on enhancing the SNR and improving speech intelligibility (Bradley et al., 1999, Bradley et al., 2003, Sato and Bradley, 2008, Roy and Browne, 2010). Listening to speech in a classroom is a complex process that not only involves the identification of the spoken word, storage, and processing of information presented but also applying it to the educational task in hand. In short, listening involves audition and cognition (Hällgren, 2005). Kahneman (1973) stresses the finite pool of cognitive resources that a listener can flexibly distribute to any task. When the primary task (listening) becomes more demanding performance on the secondary task will depreciate. The findings from the current study appear to suggest that the enhanced SNR provided by good acoustic classrooms are not enough to reduce listening effort for young people. One explanation might be that young learners require a bigger SNR provided by adaptive amplification. It appears that dynamic soundfield amplification, when combined with favourable C_{50} values, results in the listener consuming less cognitive resources when accessing the teachers' voice, which in turn provides additional cognitive resources and so positively contributes to improvements in educational outcomes.

The findings also broadly support previous studies that show young people are more adversely affected than adults when the masker is speech (Leibold and Buss, 2013). Hall et al. (2002) examined the developmental effects of perceptual masking on young people (M=7.5 years, SD=1.5) and adults (M=33.6 years, SD=9.1). A continuous meaningful speech masker composed of two competing male voices was used. The results revealed that perceptual masking was higher in young people (6.7dB) than adults (2.3dB). The authors concluded that young people may find it more difficult than adults to understand speech in natural environments when there are competing voices. The task of extracting the speech signal of interest from the noise requires increased cognitive resources and this may account for the improved scores in all the Developed Ability subtests in the classrooms exposed to dynamic soundfield.

The findings are well aligned to past research which show young people are more adversely affected by informational masking (Hall et al., 2005, Wightman and Kistler, 2005, Wightman et al., 2006, Wightman et al., 2010). One of the cues that Cherry (1953) identified which supports the release from masking is speech characteristics. If the masker and target are different then this allows the listener to separate the target signal of interest from the other competing voices. The improvements observed in the intervention classrooms may be a result of the amplified voice of the teacher being distinct from the other interfering speech maskers in the classroom. Also, it may be that the listeners' auditory attention is more focused on the amplified voice of the teacher which contributes to the release from masking. This would concur with previous research which suggests that factors that improve auditory attention can provide a release from masking (Freyman et al., 2004).

It is noteworthy that word decoding was the only reading subtest where dynamic soundfield was effective exclusively in good C_{50} classrooms. Word decoding involves the young person learning how graphemes correspond to phonemes and how they blend the phonemes into words (Schaars et al., 2017). As discussed in Chapter 2, several studies have shown that young people differ from adults in their use of acoustic information to support speech recognition. One example would be that young learners rely on formant transitions to a greater extent than adults and this process may be compromised by the presence of additional irrelevant speech sounds which masks the target signal of interest (Nittrouer and Boothroyd, 1990).

Furthermore, in suboptimal reverberant and noisy conditions, the young persons' consonant identification ability does not achieve adult-like levels of performance until the late teenage years. For tasks involving fine-grain discrimination young people require to access more acoustic energy to achieve optimal performance (Johnson, 2000). The findings from this research are well aligned with these studies. This thesis demonstrates that young people require a higher SNR when the primary task involves speech perception. This finding is interesting because it suggests that improved performance in good C_{50} classrooms cannot be attributed to good acoustics alone. Bradley et al. (2003) established that early reflection energy increases the SNR by up to 9dB and this improves speech intelligibility, especially when visual cues are removed. In classroom noise levels of 65dB SPL, the dynamic soundfield provides an SNR of ≥10dB or an ENR of ≥6dB (Phonak, 2014, Dance et al., 2018). It would appear that young learners, who do not have the same auditory perception abilities as adults, require the combination of good C_{50} values and adaptive gain to provide an appropriate level of SNR for word decoding activities.

The findings for the reading module and subtests show that dynamic soundfield was primarily effective in rooms that were classified as good and fair for speech clarity. This study was unique as it compared the intervention and control classrooms based on their C_{50} properties. No previous research has compared the efficacy of a dynamic soundfield system with a control classroom based on similar C_{50} values. Where room acoustics have been factored into previous study designs, this has been measured using the long T_{mf} of the classrooms to determine whether the rooms have positive acoustic properties for speech. Both Wilson et al. (2011a) and Dockrell and Shield (2012) found that soundfield amplification was primarily beneficial in classrooms with poor acoustics. Wilson et al. (2011a) found that student listening and auditory analysis were significantly improved in classrooms with longer RT_{60} (RT=0.087-0.091s). In the Dockrell and Shield (2012) study, soundfield amplification was also found to be effective for listening comprehension in classrooms with poor reverberation times (RT ≥0.83s). This led to the conclusion that improvements in the acoustic properties of the classroom could bring comparable benefits of soundfield amplification. However, there was no published data in either study on the C_{50} properties of the classrooms. Although it is important to avoid long reverberation times in classrooms, adding too much absorption to achieve this may be counter-productive to speech intelligibility. These results concur with Bradley et al. (2003) who reported that a room that was slightly too reverberant for speech was preferable than one that is too dead as this would lack the critical early reflection energy. The current research suggests that greater priority should be given to early reflections in classrooms.

These findings carry implications for policymakers and educators. In the word recognition, reading comprehension and standardised reading assessments the dynamic soundfield provided an educational advantage in both fair and good C_{50} classrooms. In the developed ability assessments, the dynamic soundfield was only effective in good C_{50} classrooms. Previous studies have suggested that improving the acoustic conditions of the classroom negate the requirement for soundfield systems (Dockrell and Shield, 2012). The findings from the current research would indicate that this is not the case. Even when the acoustics in the classroom provide an enhanced SNR, the learners in the dynamic soundfield classrooms demonstrated improved educational performance. One possible explanation for these findings is that young people cannot focus their attention and extract the target signal of interest, perhaps because they cannot filter out the interfering speech masker without an appropriate level of SNR.

6.10 Key findings

In relation to educational outcomes there were two primary aims of this chapter:

2) To determine if there were significant improvements in learning outcomes for Primary 3 learners (age 7 to 8) exposed to dynamic soundfield amplification in their mainstream classrooms.
3) An exploration of the relationship between the acoustic environment and dynamic soundfield on each dependent variable.

The key findings in this chapter are as follows:

- By the end of the first year in school, there was not a significant difference in reading and phonological awareness skills between learners in the control and intervention classrooms. Early mathematics demonstrated a significant difference between the classes at the end of the first year in school
- For the Developed Ability subtests and standardised module, there was a significant difference in how the control and intervention classrooms responded between the pre-intervention (baseline) and post-intervention assessments. Learners exposed to dynamic soundfield improved their picture vocabulary knowledge, non-verbal skills and developed ability skills at a significantly higher level compared to the control.
- Comparisons between the classrooms that incorporated different room conditions, indicated that for the Developed Ability module and subtests the dynamic soundfield system was primarily effective in classrooms with good C_{50} values.
- Only the General Mathematics subtests Number 2 and Data Handling showed a significant effect for the dynamic soundfield intervention. When accounting for the acoustic conditions, only Data Handling in excellent and good classrooms for speech demonstrated a significant beneficial effect.

- Mental arithmetic showed a significant time x condition interaction suggesting that there were different outcomes between the control and intervention group. The analysis showed that the classrooms exposed to dynamic soundfield had a higher change in pre-test and post-test scores compared to the control classrooms. Once again, only classrooms categorised as having good C_{50} measurements demonstrated a positive effect for the dynamic soundfield amplification.
- There were no significant improvements in spelling in the classrooms exposed to dynamic soundfield.
- All the reading subtests apart from word decoding showed a significant benefit of the intervention in classrooms with fair and good speech clarity. Significant improvements in word decoding were only observed in classrooms with good C_{50} measurements.
- Overall, subjects commonly associated with verbal communication were significantly improved in classrooms exposed to dynamic soundfield.
- Overall, subjects commonly associated with information processing were significantly improved in classrooms exposed to dynamic soundfield.

The next chapter will analyse the AfE (InCAS) results with reference to learners from the most and least deprived areas of Scotland.

Chapter 7

Results - AfE (InCAS) & SIMD

7.1 Aims

The focus of this chapter is to present the analysis of the AfE (InCAS) assessment results from SIMD 1 and SIMD 5 learners in terms of the primary research aims 2-5 of this study (see section 3.2.1). The chapter begins by discussing the statistical analysis methods used to measure any changes in the attainment gap. Thereafter, an examination of the primary 1 PIPS assessments is undertaken to provide context and to determine if there were any significant differences between the learners during the first year of school. Section 7.4 will examine the results of the AfE (InCAS) modules and subtests. Key findings are presented in Section 7.5.

7.2 Data Analysis

To answer the research questions, several two-way mixed ANOVAs were conducted on each dependent variable. Any significant interactions were followed up using paired and independent *t*-tests. The primary mixed ANOVAs used in this chapter were:

ANOVA 1	A two-way ANOVA with classroom condition (control and intervention) as the between-subject independent variable and SIMD 1 learner's AfE (InCAS) scores (pre-test and post-test) as the within-subject factor. Time (pre-test and post-test) was the repeated measure.

ANOVA 2 A two-way ANOVA with classroom condition (control and intervention) as the between-subject independent variable and SIMD 5 learner's AfE (InCAS) scores (pre-test and post-test) as the within-subject factor. Time (pre-test and post-test) was the repeated measure.

ANOVA 3 A two-way ANOVA with classroom condition (SIMD 1 learners in the intervention compared to SIMD 5 learners in the control) as the between-subject factor and one within-subject variable (pre-test and post-test). Time (pre-test and post-test) was the repeated measure.

All participants from SIMD 1 either attended classrooms categorised as fair (control, n=26, intervention, n=32) or good (control, n=41, intervention, n=56) for speech clarity. All learners from SIMD 5 attended classrooms categorised as fair (control, n=1), good (control, n=60, intervention, n=76) or excellent (control, n=19, intervention, n=27) for speech clarity. A Chi-Square test showed that the control and intervention classrooms did not significantly differ on the number of SIMD 1, χ^2 (1, N=155) =0.097, p=0.76) and SIMD 5, χ^2 (2, N=183) =0.88, p=0.65) learners in the different acoustical conditions.

To establish the effects of C_{50} on each of the dependent variables, ANOVA 1 and 2 were repeated controlling for the different acoustic conditions in the classrooms. ANOVA 3 was designed to compare the outcomes of SIMD 1 learners exposed to dynamic soundfield with SIMD 5 learners in the control. As SIMD 5 learners only attended good and excellent classrooms for speech, the inclusion of SIMD 1 learners in the fair classroom environments would introduce a confounding variable. Therefore, when comparing the outcomes between the two groups, SIMD 1 learners in the good acoustic classrooms were used for comparison.

7.3 Primary 1 PIPS standardised assessments and SIMD

Inspection of the mean and standard deviation of the early mathematics assessments reveal that learners from the most deprived quintile had an overall lower score than learners from the least deprived quintile thus indicating clear differences in educational outcomes at the point of starting school. The baseline means for SIMD 1 learners in both the control (M=89.38, SD=16) and intervention classrooms (M=92.57, SD=14) were below those from SIMD 5 learners in the control (102.15, SD=16) and intervention (106.28, SD=15) classes. Independent-*t*-tests (two-tailed) confirmed that the difference at the pre-trial stage was not significant: SIMD 1, $t(149)=-1.32$, $p=0.188$ and SIMD 5, $t(179)=-1.77$, $p=0.079$.

At the follow-up assessment, the gap between SIMD 1 and 5 learners was reduced. Two factors drove this change. Firstly, SIMD 1 learners in both the control (M=96.89, SD=18) and intervention (M=101.39, SD=15) groups demonstrated a moderate improvement in their scores from baseline. Secondly, the degree of change for SIMD 5 learners in the control (M=104.38, SD=18) and intervention groups (M=108.64, SD=13) was not as large suggesting that those from the most deprived quintiles were delayed in their development of early numeracy skills when entering school and gained more benefit from formal education. It is interesting to observe that although more progress was recorded amongst SIMD 1 learners, their end of the year mean scores were still below the level recorded at baseline by SIMD 5 learners. Independent *t*-tests confirmed the difference between the groups at the end of primary one was also not significant; SIMD 1, $t(150)=-1.64$, $p=0.102$ and SIMD 5, $t(179)=-1.86$, $p=0.064$.

The results for early mathematics align with the data from a Scottish Government (2016a) report into young learner’s development at the start of school and progress made during primary 1. In the study, it found that on entering school, learners from the most deprived quintiles were approximately thirteen months behind the expected level in early mathematics compared to learners from the least deprived quintile. By the end of the school year, SIMD 1 learners made more progress during the first year in school than SIMD 5 learners and reduced the gap in learning outcomes (the equivalent of 0.4 months development).

Looking at early reading, there was a gap between learners from the most and least deprived quintiles at the start and end of primary 1. Compared to learners from the least deprived quintile, SIMD 1 learners started school behind the expected level, and this was fairly represented in both the control (M=89.57, SD=15) and intervention classrooms (M=92.69, SD=14). The difference between the groups was not significant, $t(149)=-1.32$, p=0.188. By the end of primary 1, both the control (M=94.87, SD=15) and intervention classes (M=96.85, SD=14) showed an overall increase in their mean score, once again there was no significant difference between the groups, $t(149)=-.960$, p=0.339. It is especially interesting to note that once again learners from the least deprived quintile started school with a higher mean score than those achieved by SIMD 1 learners at the end of their first year in school.

As anticipated SIMD 5 learners had a higher mean reading score at baseline than learners from the most deprived quintile, and this was fairly represented across the control (M=100.92, SD=17) and intervention classrooms (M=103.90, SD=16). Independent *t*-tests were run on the baseline scores and this confirmed that there was no significant difference between the two groups, $t(179)=-1.19$, p=0.237. Follow-up assessments showed an improvement over time for both the control (M=106.65, SD=17) and intervention (M=109.09, SD=16) classrooms. Once, again there was not a significant difference between the two groups at the post-intervention stage, $t(179)=-0.974$, p=0.331.

Scrutiny of the change in mean scores reveals that the gap between SIMD 1 and 5 learners was larger in both the control and intervention classrooms than that observed for early mathematics. This was a result of learners from the most deprived quintile not making as large a gain over time as they did in mathematics and learners from the least deprived quintile showing a higher increase in scores. These results concur with previous research by Bradshaw (2011) which revealed that by the time of starting school, learners whose parents had no qualifications were eighteen months behind those families with a degree. One possible explanation for the improved reading scores observed in learners from SIMD 5 is a result of more support being available for early reading in the home (Scottish Government, 2016a).

Unsurprisingly, an examination of the phonological awareness means showed a gap between SIMD 1 and 5 learners at the baseline test. The results suggest that learners from the SIMD 1 quintile in the control (M=90.14, SD=15) and intervention (M=93.01, SD=15) classrooms were delayed in their skills on phonological discrimination. This is consistent with previous research by Nittrouer and Burton (2005) that found young people in areas of social deprivation have delays in perceptual strategies that support phonological processing. By the end of primary 1, both the control (M=93.29, SD=15) and intervention (M=95.53, SD=15) groups demonstrated a small improvement in scores. Independent *t*-tests, confirmed that at baseline, $t(149)=-1.12$, $p=0.265$ and follow-up, $t(149)=-0.950$, $p=0.343$ there was a non-significant difference between the groups.

Conversely, SIMD 5 learners presented with a higher mean tariff score at the baseline assessment in both the control (M=101.32, SD=14) and intervention (M=103.12, SD=13) classrooms. Changes in mean score measured at the end of primary 1 showed no major change in scores in the control (M=101.71, SD=11) and intervention classrooms (M=103.22, SD=12). These results suggest that learners from the least deprived quintile were already skilled and had knowledge of phonological awareness and so school provided very little benefit in this area. Once again there were no significant differences between the groups at baseline, $t(179)=0.824$, $p=0.411$. and follow up, $t(179)=-1.09$, $p=0.277$.

7.3.1 Summary Primary 1 PIPS standardised assessments

Examination of the PIPS assessments completed at the start and end of the first year in school identified four trends. Learners from the least deprived quintile entered school with higher test scores in all curricular areas compared to learners from the most deprived quintile. This indicates that there is already a gap in learning outcomes at the start and end of primary 1. Interestingly, learners from SIMD 5 recorded higher test scores at baseline in all three units than SIMD 1 learners achieved at the end of the first full year of schooling. This suggests that learners from the least deprived areas enter school with more of the core skills for learning. In early mathematics learners from SIMD 5 made only small gains during the school year with no noticeable changes in the phonological awareness tariff. In comparison, SIMD 1 learners made more progress in early mathematics, reducing the gap observed at baseline between SIMD 5 learners. Some small gains in phonological awareness were also recorded. This suggests, that in both assessments learners from SIMD 5 gain little benefit from school in comparison to SIMD 1 learners and this raises interesting questions about differentiation and meeting the needs of more able learners. Reading was one area in which the gap between SIMD 1 and 5 learners increased and this could reflect the additional support at home for literature in families from the least deprived quintile. These observed trends were fairly evenly represented across the control and intervention classrooms with no significant differences between them.

7.4 AfE (InCAS) assessments - Developed Ability and SIMD

7.4.1 Non-Verbal subtest

As an initial step in the statistical analysis process, comparisons between SIMD 1 and 5 learners in the intervention and control classrooms in terms of the pre-treatment subtest scores were made with independent *t*-tests to determine any significant differences prior to commencing the study. An examination of the means and standard deviation of the outcome measures showed that SIMD 5 learners in the control (M=-0.308, SD=2.1) and intervention (M=-0.346, SD=2.2) classrooms were approximately three months behind their age equivalent level prior to the study. The analysis confirmed that there was no significant difference between SIMD 5 learners in the control and intervention classrooms at the pre-trial stage, $t(181)=0.119$, $p=0.905$. SIMD 1 learners in the control (M=-1.67, SD=1.9) were approximately one year eight months behind their age equivalent level and the intervention one-year four months (M=-1.37, SD=2.3) at the pre-trial stage. Once again, there was no significant difference between the groups, $t(153)=-0.857$, $p=0.393$. Although there were no significant differences between the SIMD 1 and 5 learners in the two conditions, an attainment gap was evident between the outcomes for learners from the most and least deprived areas.

	Within-Subject			Between-Subject			Interactions		
α=0.05	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
ANOVA 1 (1,153)	31.93	0.001	0.173	4.60	0.034	0.029	4.43	0.037	0.028
ANOVA 2 (1,181)	42.67	0.001	0.191	0.128	0.721	0.001	0.943	0.333	0.005
ANOVA 3 (1,133)	55.72	0.001	0.295	4.31	0.040	0.031	5.18	0.025	0.037

Table 48: Results from the two-way mixed ANOVAs conducted on the non-verbal ability subtest.

The results for ANOVA 1 are presented in Table 48 and reveal a large and significant increase in AfE (InCAS) scores between the pre-test and post-test assessments. Applying the Bonferroni correction ($\alpha/4 = 0.0125$), further analysis confirmed that the control classrooms made non-significant progress during the study, $t(66)=-2.34$, $p=0.022$. In contrast, the intervention groups, $t(87)=-5.92$, $p<0.001$ made significant progress during the academic year. The magnitude of the progress was greater in the intervention classrooms (M=-1.34, SD=2.1) compared to the control (M=-0.613, SD=2.1). There was also a significant main effect of condition, indicating a significant difference between the treatment and comparison groups. The condition x time interaction was also significant showing that one group changed their scores more than the other between the pre-test and follow-up assessments. Figure 56 demonstrates the nature of the interaction, showing that the intervention group (green line) made the greatest progress between the pre-test and post-test scores.

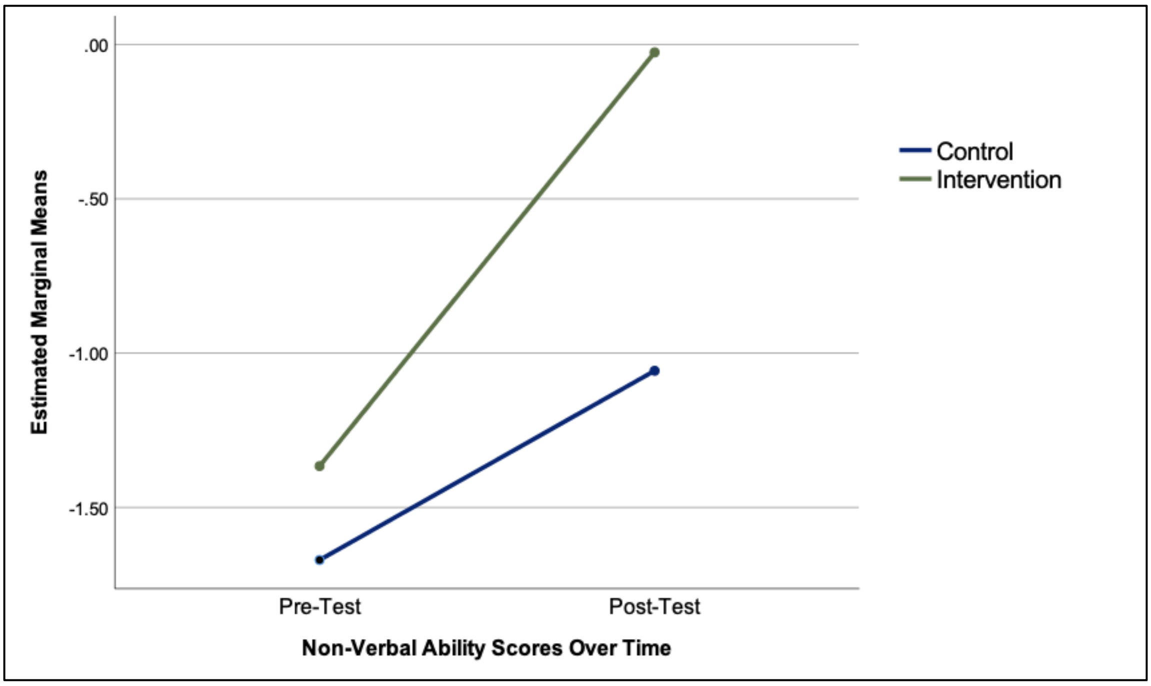


Figure 56: Interactional line chart showing the change in pre-test and post-test scores in the control (blue line) and intervention (green line) classrooms for SIMD 1 learners.

Follow-up independent *t*-tests (two-tailed) decomposed the interaction further to establish if there were significant differences between the groups at the end of the intervention. Unlike at the pre-intervention stage, there was a significant difference at the end of the study between both groups, $t(153)=-2.89$, $p=0.004$. The magnitude of the difference between the two groups was moderate ($d=0.49$). Although both the control and intervention classrooms demonstrated progress over the academic year, the difference between the groups was only evident for the post-intervention scores. Overall, the results suggest that in comparison to the control classrooms, learners from the most deprived quintile exposed to dynamic soundfield showed significantly greater improvements in non-verbal ability skills.

In contrast, ANOVA 2 with SIMD 5 learners only showed a significant main effect of time with no condition or interaction effect. As there were no significant differences between the groups and no significant interactions then any difference in the pre-test and post-test scores cannot be attributed to the dynamic soundfield intervention, as such no further analysis was run.

To explore the relationship between the different acoustical settings and the outcomes observed, ANOVA 1 was repeated to determine if there was an effect for rooms with different classifications of speech clarity. Only classrooms that were categorised as good for speech weighted C_{50} demonstrated a significant interaction was present. In the good classrooms for speech, there was a significant main effect for the change of scores over time, $F(1, 95)=27.63$, $p<0.001$, $\eta_p^2=0.225$ and an interaction effect of time and condition, $F(1, 95)=5.84$, $p=0.017$, $\eta_p^2=0.058$. Figure 57 graphically illustrates the interaction.

Paired *t*-tests showed that the control group made no significant progress from time point 1 and time point 2, $t(40)=-1.98$, $p=0.060$. In contrast, the intervention group showed a significant improvement in test scores, $t(55)=-5.70$, $p<0.001$. Independent *t*-tests were run on the pre-intervention and post-intervention scores to establish if there were any between group differences. A non-significant difference was observed at the pre-test point, $t(95)=-0.069$, $p=0.945$ and a significant difference after the experimental phase, $t(95)=-2.73$, $p=0.007$. The effect size for the difference between the two groups was medium ($\boldsymbol{d}=0.50$). Overall, SIMD 1 learners exposed to dynamic soundfield in rooms good for speech clarity demonstrated a significant improvement in outcomes in non-verbal ability. In contrast, the learners in classrooms that were categorised as fair for speech weighted C_{50} showed no significant effect for the intervention. The results suggest that the effectiveness of the dynamic soundfield intervention for non-verbal ability is influenced by the acoustical properties of the classroom.

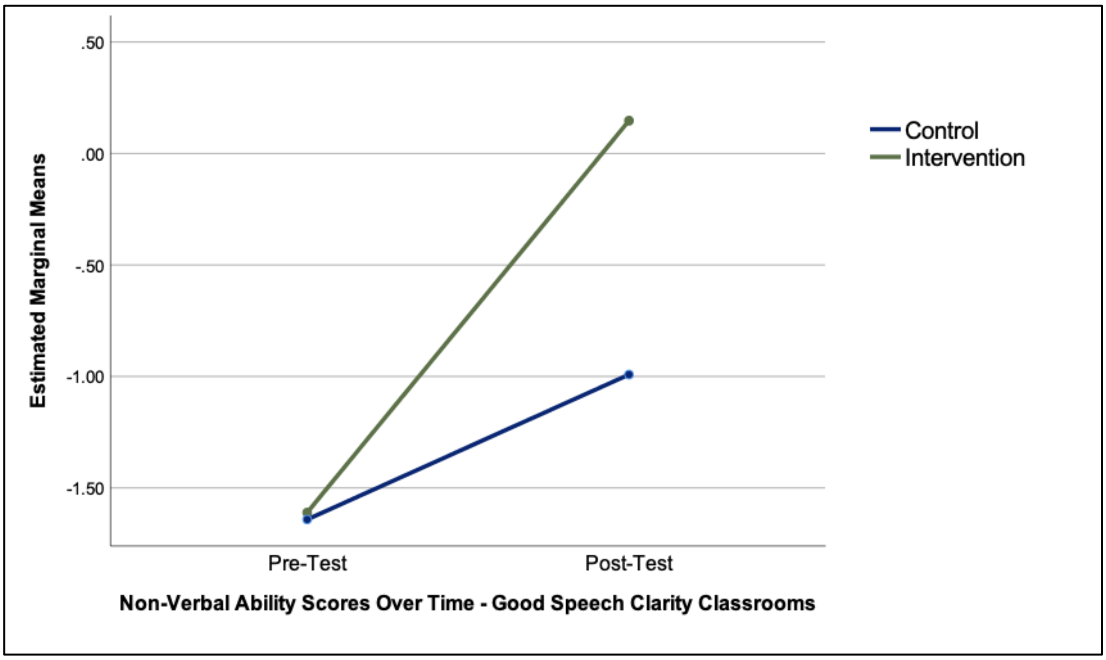


Figure 57: Interactional line chart showing the change in non-verbal pre-test and post-test scores in the control (blue line) and intervention (green line) classrooms for SIMD 1 learners in good speech weighted C_{50} classrooms (n=18).

ANOVA 3 compared SIMD 1 learners in the intervention classrooms (good speech clarity classrooms) with SIMD 5 learners in the control condition (good and excellent classroom environment). As Table 48 shows there was a significant main effect of time, condition and condition x time interaction. As Figure 58 illustrates, it appears changes in the post-test scores of the SIMD 1 intervention group drove the interaction. As is evident, there were significant differences between the two groups at baseline, with SIMD 1 learners demonstrating greater difficulty than SIMD 5 learners with non-verbal reasoning skills, $t(133)=-2.64$, p=0.009. The analysis confirmed that both the SIMD 1 learners in the intervention classrooms, $t(55)=-5.23$, p<0.001 and SIMD 5 learners in the control group, $t(78)=-4.97$, p<0.001 made significant progress during the academic year. The progress was greater for SIMD 1 learners (M=-1.72, SD=2.4) compared to SIMD 5 (M=-.917, SD=1.6). By the end of the study, there was a non-significant difference between the groups,

$t(133)=0.713$, p=0.477. Interestingly, the gap between learners from the most and least deprived quintiles was reduced on the post-test scores and this was the result of a larger change in scores recorded in learners exposed to dynamic soundfield.

The progress made by SIMD 1 learners, is partially explained, but not fully, by maturation since they started the year significantly delayed in non-verbal ability compared to SIMD 5 learners. Both the control (M=-0.613, SD=2.1) and intervention (M=-1.34, SD=2.1) groups demonstrated a change in scores over the course of the study. However, the significant differences between SIMD 1 learners shown in ANOVA 1 and the interaction detected in Figure 58 suggest that the dynamic soundfield intervention contributed to the improved scores.

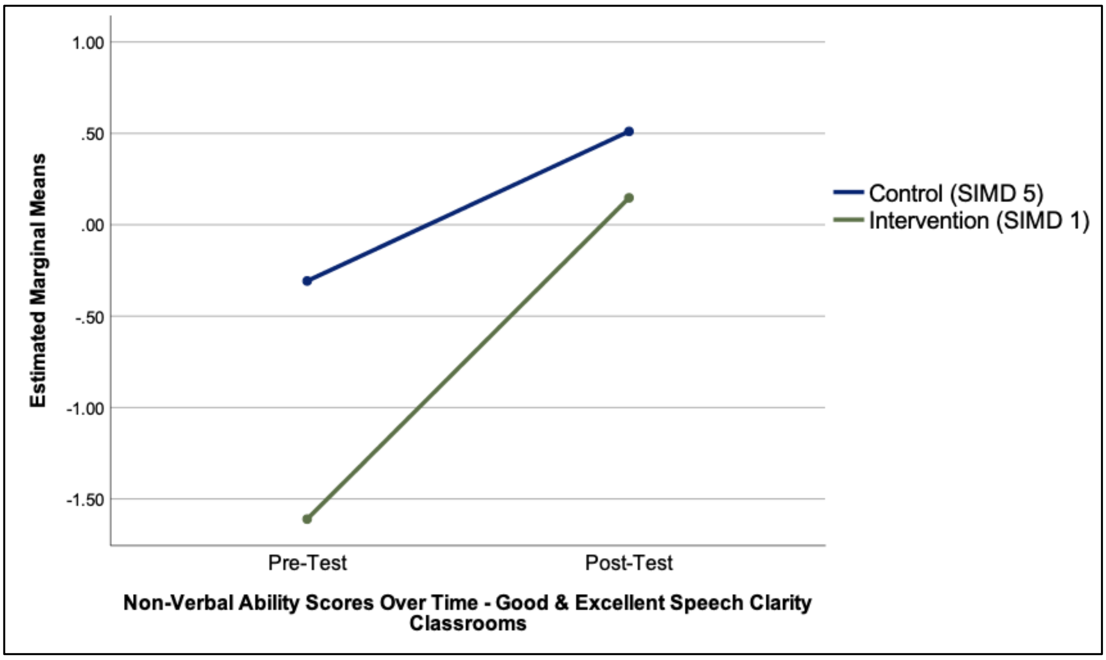


Figure 58: Interactional line chart showing the change in non-verbal subtest scores from the pre-test and post-test assessments.

7.4.2 Picture Vocabulary subtest

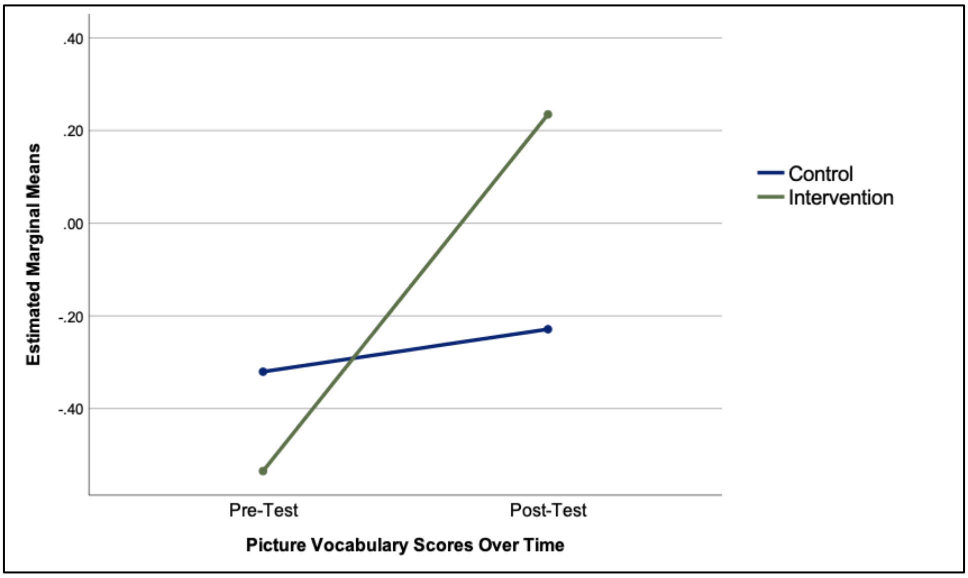
Analysis of pre-intervention scores reveals that there was no significant difference between SIMD 1 learners in the control and dynamic soundfield classrooms before the intervention, $t(153)=-0.318$, $p=0.361$. The control groups were approximately 3 months behind their age equivalent stage (M=-0.320, SD=1.5) and the intervention classroom 6 months behind (M=-0.535, SD=1.4). In contrast, SIMD 5 learners were approximately eight months ahead of their age equivalent stage in both the control (M=0.706, SD=1.8) and intervention (M=0.760, SD=1.7) classrooms. Unsurprisingly, the analysis confirmed there was a non-significant difference between SIMD 5 learners in the control and intervention classrooms, $t(181)=-0.204$, $p=0.839$. As with the non-verbal ability scores, an attainment gap was evident between learners from the most and least deprived quintiles.

	Within-Subject			**Between-Subject**			**Interactions**		
α=.05	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
ANOVA 1 (1,153)	14.09	0.001	0.084	0.301	0.584	0.002	8.73	0.004	0.054
ANOVA 2 (1,181)	13.55	0.001	0.070	0.399	0.528	0.005	0.983	0.323	0.005
ANOVA 3 (1,133)	22.39	0.001	0.144	11.79	0.001	0.081	5.76	0.018	0.042

Table 49: Results from the two-way mixed ANOVAs conducted on the picture vocabulary subtest.

Looking firstly at ANOVA 1, the results in Table 49 reveal a large and significant effect of time, a non-significant effect of condition and a significant condition x time interaction. As Figure 59A graphically illustrates, the crossed blue (control) and green (intervention) lines show that a disordinal interaction is present. Once again, as the main effects cannot be reliably interpreted the interactions were instead investigated with independent and paired *t*-tests. The analysis revealed that the control classrooms, $t(66)=-0.490$, $p=0.626$ made no significant progress during the year. In contrast, the classrooms fitted with dynamic soundfield made significant progress, $t(87)=-5.47$, $p<0.001$. This is evident in the gradient of the control and intervention lines in Figure 59A. Analysis of the mean scores showed a significant difference at the end of the experimental stage, $t(153)=-2.89$, $p=0.004$ suggesting, as with the Non-Verbal Ability subtest, that learners from the most deprived areas gain a significant benefit of dynamic soundfield amplification.

A)



B)

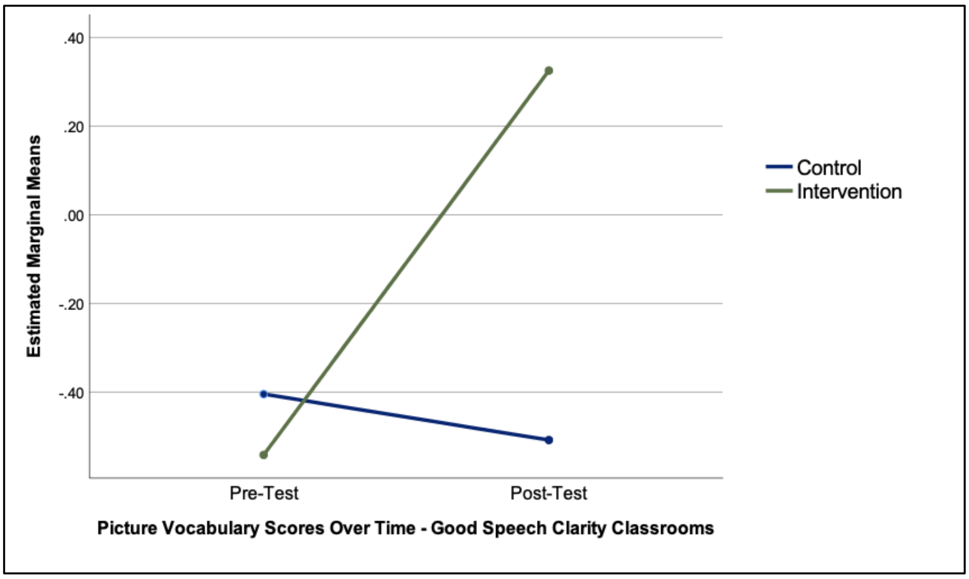


Figure 59: A) Disordinal interaction between time and condition for the Picture Vocabulary subtest for SIMD 1 learners. B) ANOVA repeated in classrooms with good speech clarity (n=18).

Further analysis was performed to establish the effects of classroom acoustics on the observed outcomes. A mixed ANOVA revealed that in good classrooms for speech there was a main effect of time, $F(1, 95)=6.36$, $p=0.013$, $\eta_p^2=0.063$ and a significant interaction between time and condition, $F(1, 95)=10.28$, $p=0.002$,

$\eta_p^2=0.098$ indicating there were significant differences in scores between the two groups. The fair classrooms demonstrated no significant interactions or between-group differences; therefore, no further analysis was performed.

Inspection of Figure 59B shows that the change in results in the good C_{50} classrooms was caused by a slight decrease in scores in the control group (blue line) and a larger increase in scores in the intervention group (green line). Although both groups were exposed to a classroom environment that had positive early reflections for speech, it was only SIMD 1 learners in the intervention classrooms that demonstrated significant progress. Two trends are apparent. Firstly, the dynamic soundfield system is primarily beneficial in rooms with good C_{50} properties. The second is that good acoustics alone do not appear to contribute to improved performance in picture vocabulary knowledge for SIMD 1 learners. Follow up tests on the disordinal interaction confirmed this; decreases in the control were non-significant, $t(40)=-0.390$, $p=0.698$, in contrast, the intervention increases were significant, $t(55)=-5.06$, $p<0.001$. Looking at the change of scores at the pre-intervention stage showed there was no difference between the groups, $t(95)=0.404$, $p=0.661$ and a significant difference at the end of the experimental phase, $t(95)=-2.35$, $p=0.021$. In keeping with the previous results, this suggests that dynamic soundfield was primarily effective in good acoustic environments.

ANOVA 2 also showed a large and significant main effect of time. Paired *t*-tests (two-tailed) showed that the control, $t(78)=-1.72$, $p=0.089$ group made no significant progress between the time of assessment 1 and assessment 2. In contrast, the intervention classrooms made significant progress, $t(103)=-3.66$, $p<0.001$. There was a non-significant effect of condition and interaction. The results again suggest that for learners in the least deprived areas there is no significant benefit gained from dynamic soundfield amplification.

As Table 49 shows, ANOVA 3 revealed a significant main effect of time and condition. Furthermore, there was a significant interaction indicating that the change in scores was greater in one group compared to the other over time. Analysis of Figure 60 shows that the gradient of the intervention line (green) is steeper than the control (blue) suggesting that the intervention group drove this interaction. Follow-up tests on the significance of the interaction showed that SIMD 1 learners exposed to dynamic soundfield made significant progress between time points 1 and 2, $t(55)=-5.05$, $p<0.001$. SIMD 5 learners in the control demonstrated a non-significant effect of time, $t(78)=-1.72$, $p=0.089$. The results suggest that for the Picture Vocabulary subtest there was a significant difference in how SIMD 1 learners exposed to dynamic soundfield responded over the time of the study compared to SIMD 5 learners in the control. These combined factors appear to have affected the outcome observed. Between-subject analysis unsurprisingly showed that prior to the intervention (MD=1.25) there was a significant difference between learners from SIMD 1 and SIMD 5, $t(133)=-4.21$, $p<0.001$. By the end of the intervention (MD=0.66), there remained a significant difference, $t(133)=-2.13$, $p=0.035$. Overall, the results suggest that SIMD 1 learners in the intervention classrooms with good speech gained the greatest benefit from dynamic soundfield amplification. Although the gap between the most and least deprived quintiles was reduced, this was not statistically significant.

Figure 60: Change in the picture vocabulary pre-test and post-test scores for SIMD 1 learners in the intervention classrooms (green line) and SIMD 5 learners in the control.

7.4.3 Developed Ability module

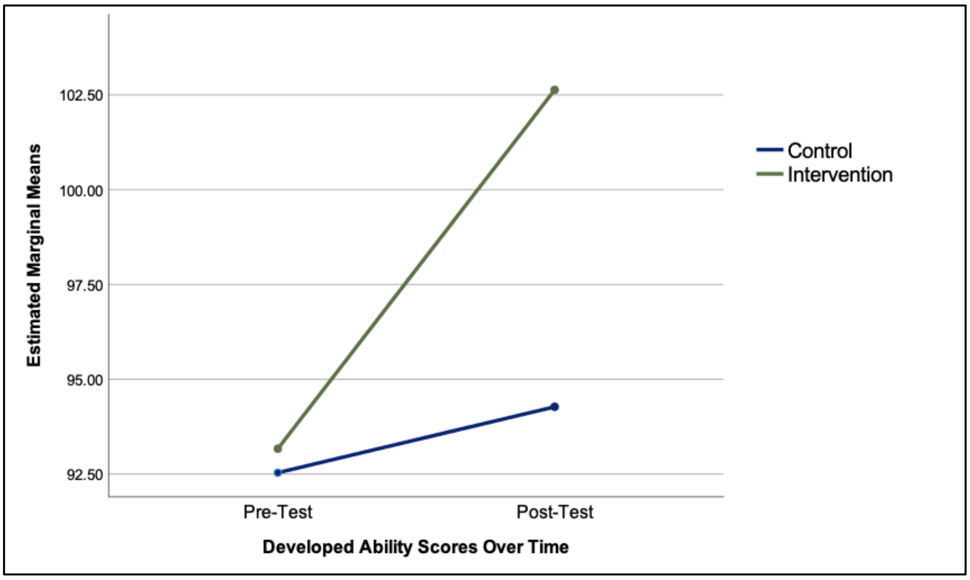
Analysis of the pre-treatment standardised results suggests that there were no significant differences in developed ability levels between learners from the most deprived quintile in the control and intervention groups, $t(153)=-0.321$, $p=0.748$. For SIMD 1 learners in the control (M=92.53, SD=12.2) and intervention (M=93.17, SD=12.1) classes, the standardised scores were fairly even matched prior to the intervention. SIMD 5 learners in both the control (M=103.70, SD=15.1) and intervention classrooms (M=103.84, SD=15) had an overall higher mean score than SIMD 1 learners. The difference between SIMD 5 learners in the control and intervention classrooms prior to the intervention were non-significant, $t(181)=-0.066$, $p=0.947$.

	Within-Subject			**Between-Subject**			**Interactions**		
α=0.05	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
ANOVA 1 (1,153)	30.99	0.001	0.168	5.67	0.018	0.036	14.74	0.001	0.088
ANOVA 2 (1,181)	51.23	0.001	0.221	0.152	0.697	0.001	0.723	0.396	0.004
ANOVA 3 (1,133)	70.28	0.001	0.346	12.86	0.001	0.089	7.96	0.006	0.056

Table 50: Results from the two-way mixed ANOVAs conducted on the standardised Developed Ability module.

As Table 50 shows, results from ANOVA 1 revealed a significant main effect of the change in scores over time. As Figure 61A illustrates a significant interaction was observed with the intervention group (green line) scores changing at a higher level between the two test periods compared to the control (blue line). Post-intervention assessments were significantly higher than at baseline. To answer the research question of whether there was a significant difference in scores between SIMD 1 learners in the control and intervention classrooms follow-up tests on the interaction were performed. Changes in pre-test/post-test scores for the control groups were found to be non-significant, $t(67)=-1.04$, $p=0.301$. In contrast, the classrooms exposed to dynamic soundfield significantly improved their overall scores, $t(88)=7.80$, $p<0.001$. The results show there is a significant improvement in developed ability scores only for learners that attended classrooms fitted with dynamic soundfield. Independent *t*-tests (two-tailed) also show that when the same analysis was conducted at the end of the intervention there was a significant difference between the two groups, $t(153)=-3.64$, $p<0.001$. The combined analysis demonstrates that learners from the most deprived quintile made significant progress in their developed ability knowledge compared to SIMD 1 learners in the control.

A)



B)

Figure 61: A) Interaction between time and condition for the Developed Ability module for SIMD 1 learners. B) ANOVA repeated in classrooms with good speech clarity (n=18).

ANOVA 1 was repeated controlling for the acoustic conditions. In good classrooms for speech a main effect of time, $F(1, 95)=17.36$, $p<0.001$, $\eta_p^2=.155$, condition, $F(1, 95)=6.27$, $p=0.014$, $\eta_p^2=0.062$ and interaction was observed, $F(1, 95)=12.43$, $p<0.001$, $\eta_p^2=0.116$. As Figure 61B shows, the modest elevation of the blue line (control group) in comparison to the steep gradient of the green line (intervention group) indicates the magnitude of change in scores was higher in the intervention classrooms. Follow-up results confirmed that only the dynamic soundfield classrooms showed a significant increase in scores, $t(55)=-6.76$, $p<0.001$. Between-group analysis also confirmed that there was only a significant difference between the groups at the follow-up stage, $t(95)=-3.63$, $p<0.001$. The magnitude of the difference between the two groups was large ($\boldsymbol{d}=0.72$). The results suggest that in good acoustic environments for speech learners from the most deprived quintile made moderate progress during the school year, in contrast, learners exposed to dynamic soundfield showed a significant improvement. Once again, fair classrooms did not achieve significance. The results once again suggest that good acoustics alone cannot provide the beneficial effects observed when the listening environment is supplemented by adaptive amplification.

ANOVA 2 was undertaken to establish if there were a differential effect of dynamic soundfield for SIMD 5 learners in the control and intervention classrooms. Unsurprisingly based on the subtest analysis, there was only a main effect for time and no effect of condition and interaction. Due to the results, the assessments were not repeated. Overall, the results suggest that SIMD 5 learners in the intervention classrooms showed no significant effect of the dynamic soundfield intervention.

ANOVA 3 revealed that, in comparison to the control group, the intervention classrooms showed significantly greater improvements in their scores between the pre-test and post-test stages. As Table 50 reveals there was also a significant main effect of condition, with the control group (SIMD 5) presenting with a higher mean score ($M=108.71$, $SD=15.9$) at the end of the treatment compared to the intervention

(SIMD 1) (M=103.23, SD=10.7). Figure 62 graphically illustrates the condition x time interaction, showing the differential improvement in the control (blue line) and intervention (green line) classrooms. There was a significant difference between the two groups prior to the intervention, $t(133)=-4.46$, $p<0.001$ and at the end of the study, $t(133)=-2.25$, $p=0.026$. Analysis of the F and p values show there was a reduction in the gap between the SIMD 1 learners in the intervention and SIMD 5 learners in the control. Although there was a reduction in the gap between the two groups, this was not significant.

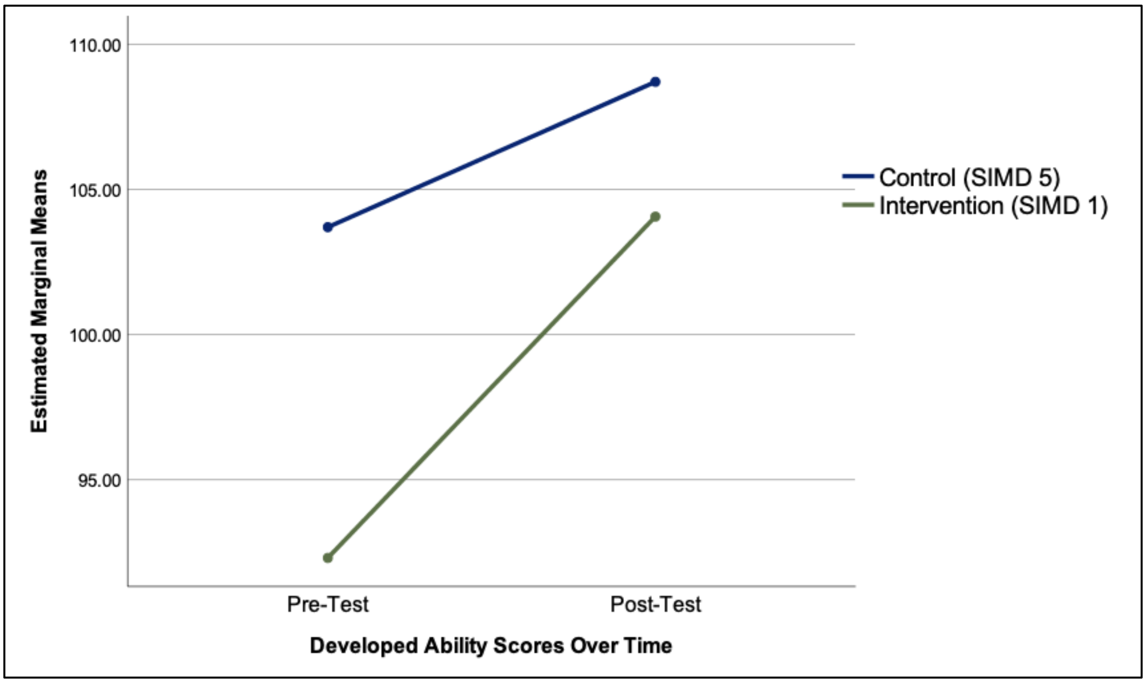


Figure 62: Interactional line chart showing the change in standardised Developed Ability module scores between SIMD 1 intervention learners (green line) and SIMD 5 control (blue line).

7.4.4 Summary Developed Ability module and subtests

The results from the Developed Ability module and subtests identified two trends. Learners from SIMD 1 in the intervention classroom gained a significant benefit from dynamic soundfield, in comparison to the SIMD 1 control group. SIMD 5 in both the control and intervention classrooms showed no significant effect for the dynamic soundfield system. Furthermore, the dynamic soundfield system was mainly effective in classrooms that had good C_{50} properties for speech. Previous research has indicated that young people in classrooms exposed to noise are more likely to disengage from tasks and have a higher failure rate (Cohen et al., 1980). The results from the current study suggest that young people that are significantly delayed in their cognitive development gain an advantage from a dynamic soundfield system. One reason may be that by making the voice of the teacher distinct from the competing voices in the room it is easier for the listener to identify the target signal of interest in a multi-talker environment. This reduces the amount of listening effort required which is beneficial to the rapid and efficient processing of information.

7.5 AfE (InCAS) assessments - General Mathematics and SIMD

7.5.1 Number 1 and Number 2 subtests

Concerning the pre-intervention Number 1 subtest, results from the independent *t*-tests revealed a non-significant difference between learners from SIMD 1 in the control (M=-0.444, SD=1.2) and treatment groups (M=-0.311, SD=1), $t(153)=-0.751$, $p=0.454$. Analysis of the Number 2 subtest on the same cohort, also found that there was not a significant difference, $t(153)=-1.57$, $p=0.118$ between the control (M=-0.429, SD=1.1) and intervention groups (M=-0.137, SD=1.1). Learners from the least deprived quintile also had a non-significant difference between the groups at the pre-intervention stage in the Number 1 subtest, $t(181)=-0.113$, $p=0.910$. Examination of the mean and standard deviation shows that SIMD 5 learners, in both the control (M=0.064, SD=1.1) and intervention classrooms (M=0.044, SD=1.2) are marginally ahead of their age-appropriate level, in contrast, the SIMD 1 learners in the control group were approximately 5 months behind their chronological stage and the intervention classes 3 months behind. Analysis of the Number 2 assessment also revealed there was no significant difference, $t(181)=-0.699$ $p=0.485$ between the groups of SIMD 5 learners. SIMD 5 learners in the control (M=0.006, SD=1.2) were approximately at an age-appropriate level and the intervention (M=0.146, SD=1.4) classrooms were approximately one month ahead.

ANOVA 1 was run on the Number 1 subtest and found no significant main effects of time, $F(1, 153)=0.186$, $p=0.667$, $\eta_p^2=0.001$, condition, $F(1, 153)=1.95$, $p=0.165$, $\eta_p^2=0.013$ or interaction, $F(1, 153)=3.01$, $p=0.085$, $\eta_p^2=0.002$. Based on these results, it was unnecessary to repeat the ANOVA controlling for the good and fair acoustical conditions. On the Number 2 subtest, results revealed a significant main effect of time, $F(1, 153)=4.91$, $p=.0028$, $\eta_p^2=0.031$. There was also a significant main effect of condition, $F(1, 153)=5.17$, $p=0.024$, $\eta_p^2=0.033$. The post-test scores were found to be significantly different between the control and intervention classrooms, $t(153)=-2.24$, $p=0.027$. However, there was no time by condition interaction. ANOVA 2 and ANOVA 3 also found there to be no significant interaction, condition or main effects. Overall, the results from the three ANOVAs suggest that the intervention had no significant effect on the outcomes observed.

7.5.2 Data Handling subtest

To ensure there was parity between learners from SIMD 1 and 5 in the two groups an independent *t*-test was performed on the baseline assessments. No significant differences were found between learners from the least deprived quintiles, $t(181)=0.699$ $p=0.485$. It is noteworthy that the means scores in both the control (M=0.191, SD=1.4) and intervention (M=0.270, SD=1.3) classrooms were slightly higher than for the Number 1 and 2 subtests. However, SIMD 1 learners in both the control (M=-0.609, SD=1) and intervention (M=-0.303, SD=1) classrooms were behind their age-appropriate level. The difference between the groups was not significant, $t(153)=-1.87$ $p=0.064$

ANOVA 1 revealed a significant main effect of time, $F(1, 153)=14.28$, $p<0.001$, $\eta_p^2=0.083$. Age difference scores over the time of the study were found to be significantly improved only in the intervention group, $t(103)=-4.46$, $p<0.001$. A significant effect of condition was also revealed, $F(1, 153)=4.71$, $p=0.031$, $\eta_p^2=0.030$. The post-test scores demonstrated a significant difference in outcomes between the intervention and control group, $t(103)=-2.52$, $p=0.026$. However, more importantly, the interaction between time and condition was not significant therefore indicating that the changes in the dependent variable did not appear to be connected to the intervention. The results were repeated to control for the acoustic conditions with both fair and good classrooms for speech only showing a significant main effect of time. The overall results suggest that dynamic soundfield did not have a differential effect on learners from the most deprived areas.

In ANOVA 2, the test of within-subject revealed a significant difference between the pre-test and post-test scores, $F(1, 181)=10.30$, $p=0.002$, $\eta_p^2=0.054$. Only the classrooms exposed to dynamic soundfield demonstrated a significant improvement in scores between the two assessment time points, $t(103)=-4.97$, $p<0.001$. Mean difference values indicate that the control group ($MD=-0.152$, $SD=1.4$) recorded a modest change in scores between the pre-test and post-test scores, in comparison the intervention classroom ($M=-0.556$, $SD=1.1$) scores were approximately 27.3% higher. Concerning the interaction, significance was also achieved, $F(1, 181)=4.74$, $p=0.031$, $\eta_p^2=.026$. As Figure 63 illustrates, the post-test scores were not only better in the intervention classrooms compared to the control, but there was also a significant difference between the groups at the end of the study, $t(181)=-2.39$, $p=0.018$. This demonstrates that the changes in post-test scores for data handling were higher for SIMD 5 learners in the intervention classrooms compared to the control cohort.

Figure 63: Interactional line chart showing the change in Data handling scores between SIMD 5 learners in the intervention (green line) and control group (blue line).

When examining the effects of the good classroom environment for C_{50} on the data handling outcome, there was a significant effect of time, $F(1, 134)=13.41$, $p<0.001$, $\eta_p^2=0.091$ and condition, $F(1, 134)=3.95$, $p=0.049$, $\eta_p^2=0.029$ but no significant interaction, $F(1, 134)=1.75$, $p=0.189$, $\eta_p^2=0.013$. As the interaction between condition and time did not achieve significance, no further analysis was run.

In the excellent classrooms for speech, there was a significant disordinal interaction, $F(1, 44)=4.69$, $p=0.036$, $\eta_p^2=0.096$. As Figure 64 illustrates, the factors driving the interaction are the decrease in scores in the control classroom that was not significant, $t(18)=-0.590$, $p=0.60$, and a steep significant change in scores in the intervention classrooms, $t(26)=-2.68$, $p=0.013$. The results indicate that the changes in data handling scores may be attributable to the beneficial effect of the dynamic soundfield in classrooms with excellent C_{50} values. However, these findings should be treated with caution, particularly as the sample size was small and only two classrooms were involved.

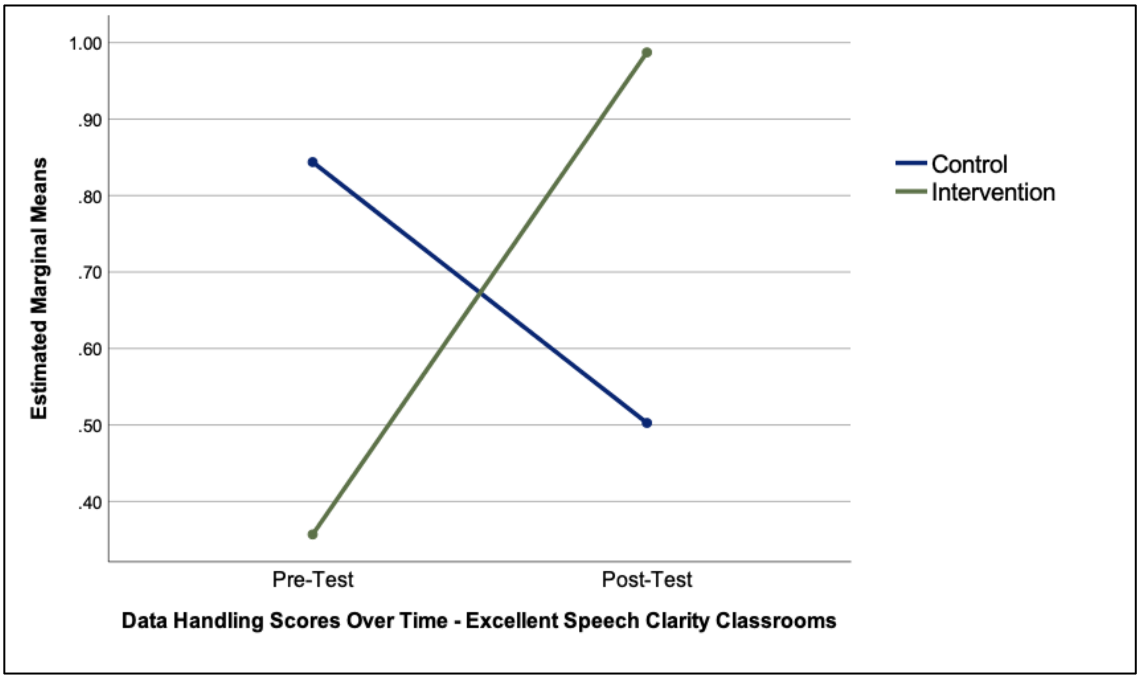


Figure 64: Interactional line chart showing the change in Data handling scores between SIMD 5 learners in the intervention (green line) and control group (blue line) in classrooms with excellent C_{50} values (n=2).

ANOVA 3 found no significant main effect of condition, $F(1, 133)=2.42$, $p=0.122$, $\eta_p^2=0.01$ or interaction, $F(1, 133)=.385$, $p=0.60$, $\eta_p^2=0.003$. The two-way mixed ANOVA did reveal a significant effect of time, $F(1, 133)=3.94$, $p=0.049$, $\eta_p^2=0.029$. As no significant interaction was found between the pre-test and post-test scores and there was no association between the classrooms the participants attended, no further analysis was carried out.

7.5.3 Measure, Shape and Space subtest

Inspection of the means reveals, that once again SIMD 1 learners in the control (M=-0.366, SD=1.2) and intervention groups (M=-0.135, SD=1), at baseline are behind their age appropriate level. The difference between the groups was not significant, *t*(153)=-1.30, p=0.196. The highest baseline scores were again found with SIMD 5 learners in both the control (M=0.158, SD=1.1) and intervention groups (M=0.514, SD=1.1). The analysis revealed a significant difference between the groups, before the experimental phase, *t*(181)=-2.14, p=0.034.

	Within-Subject			Between-Subject			Interactions		
α=0.05	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
ANOVA 1 (1,153)	3.61	0.059	0.023	1.13	0.290	0.007	0.481	0.489	0.003
ANOVA 2 (1,181)	11.28	0.001	0.059	4.69	0.032	0.025	0.005	0.943	0.001
ANOVA 3 (1,133)	4.88	0.029	0.035	8.21	0.005	0.058	0.333	0.565	0.002

Table 51: Results from the two-way mixed ANOVAs conducted on the subtest Measure, Shape and Space.

As Table 51 illustrates ANOVA 1 found no main effects or interactions. Based on these results, ANOVA 1 was not repeated for the different acoustical conditions. Both ANOVA 2 and 3 identified significant main effects of time and condition. Importantly, there was no significant interaction, indicating that whether you attended the control or intervention classrooms did not have a moderating effect on the dependent variable. As there were main effects, both ANOVAs were repeated taking into account the C_{50} values of the classrooms.

ANOVA 2 found that in the excellent classrooms for speech there was a significant disordinal interaction effect, $F(1, 181)=4.66$, $p=0.036$, $\eta_p^2=0.096$. Figure 65 illustrates that the intervention group showed a greater amount of change than the control group over the two-time points. Paired *t*-tests confirmed that only the intervention group demonstrated a significant improvement over time, $t(26)=-3.98$, $p<0.001$. Once again, these findings should be treated with caution, as the sample size was small and only two classrooms were involved. There were no significant effects for the classroom with good C_{50} values. ANOVA 3 found no significant interactions in excellent and good classroom environments. Overall, the results suggest that only learners from the least deprived quintile attending classrooms with excellent speech clarity gained a significant benefit from dynamic soundfield amplification.

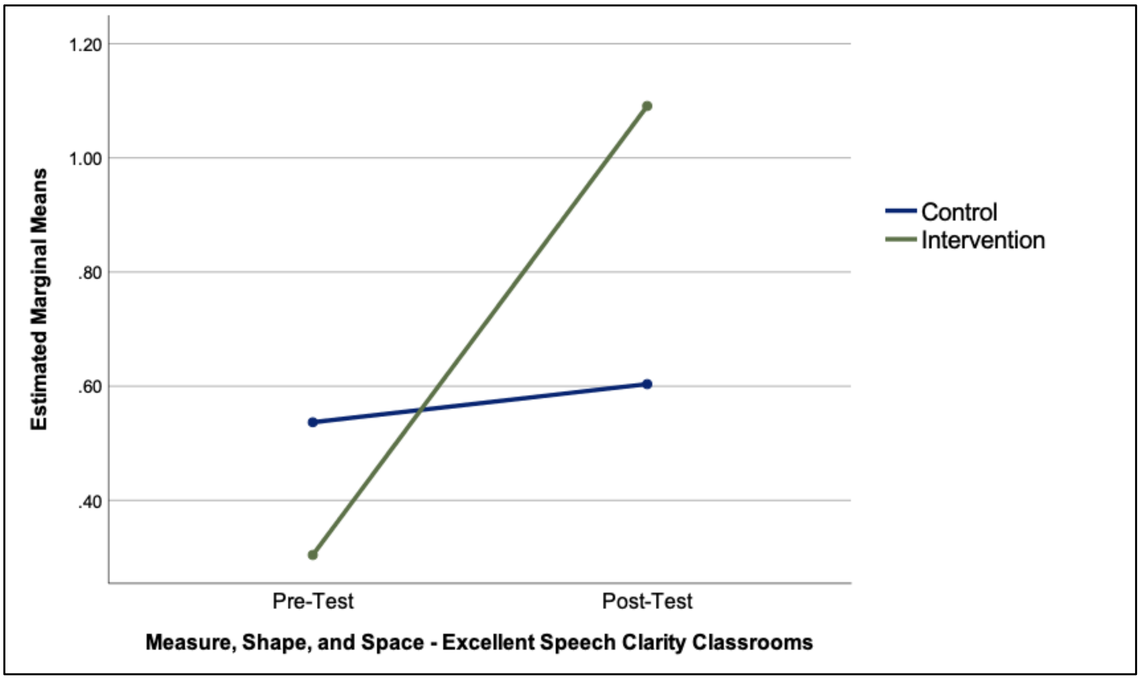


Figure 65: Interactional line chart showing the change in measure, shape and space subtest from the pre-test and post-test assessments for SIMD 5 learners in excellent C_{50} conditions (n=2).

7.5.4 General Mathematics module

Inspections of the means and standard deviation reveal that SIMD 5 learners achieved a higher baseline standardised score than SIMD 1 learners. Pre-treatment, there was a significant difference, $t(181)=-2.05$, $p=0.042$ between SIMD 5 learners in the control (M=103.15, SD=14) and intervention classrooms (M=107.40, SD=14). There was also a significant difference, $t(153)=-2.52$, $p=0.013$ observed for SIMD 1 learners in the control (M=94.12, SD=12) and intervention groups (M=99.11, SD=12).

ANOVA 1 demonstrated significant main effects for time, $F(1, 153)=4.66$, $p=0.036$, $\eta_p^2=0.096$ and condition, $F(1, 153)=5.73$, $p=0.018$, $\eta_p^2=0.036$. When the assessment was repeated at the end of the study inspection of the means indicated that neither group showed a significant difference in scores. As no interaction was found, $F(1, 153)=0.596$, $p=0.478$, no further analysis was conducted. ANOVA 2 also followed a similar pattern, only finding a significant effect of time, $F(1, 181)=20.65$, $p<0.001$, $\eta_p^2=0.102$ and condition, $F(1, 181)=5.37$, $p=0.022$, $\eta_p^2=0.029$. Once again, no interaction was found, $F(1, 181)=1.06$, $p=0.305$, $\eta_p^2=0.006$ and so no further analysis was conducted. ANOVA 3 found no main effects and no interactions.

7.5.5 Summary – General Mathematics module and subtests

The efficacy of dynamic soundfield at reducing the attainment gap in general mathematics was addressed in this section. The results from the mixed ANOVAs comparing SIMD 1 learners in the intervention classroom with SIMD 5 learners in the control found no evidence to support a reduction in the poverty-associated attainment gap between the two quintiles. When investigating whether there was an improvement in scores for SIMD 1 and 5 learners exposed to dynamic soundfield, in comparison to their counterparts in the control condition, a significant interaction was only found for SIMD 5 learners in data handling and measure, shape and space. The data handling subtest involves locating and interpreting data in a table, extracting and evaluating information. Measure, Shape, and Space involve simple coordinates, comparing and evaluating symmetrical shapes. Scrutiny of the results shows that the intervention was primarily effective in classrooms that were excellent for speech clarity. As described previously, the small sample size and the limited number of classrooms means that the results need to be treated with caution.

7.6 AfE (InCAS) assessments -Mental Arithmetic and SIMD

With reference to SIMD 1 learners, an inspection of the means in both the control (M=-1.34, SD=1.4) and intervention (M=-1.02, SD=1.5) classrooms revealed there was an age equivalent deficit of over one year. Although the age equivalent difference was higher in the control group, this was not statistically significant, $t(153)=-1.33$, p=0.186. In contrast, SIMD 5 learners were behind their age equivalent level by approximately 3 months in the control (M=-0.341, SD=1.6) and 1 month (M=-0.123, SD=1.6) in the intervention classrooms. Once again the difference between the groups was not significant, $t(181)=-0.930$, p=0.353. These findings suggest that both SIMD 1 and 5 learners were behind their expected stages in mental arithmetic. When compared the mean scores show that SIMD 1 demonstrated the largest deficit which once again illustrates a poverty-associated attainment gap.

ANOVA 1 identified a significant effect of time, $F(1, 153)=9.04$, $p=0.003$, $\eta_p^2=0.056$ suggesting there was a significant difference between the pre-test and post-test scores. Concerning the between-subjects test, $F(1, 153)=1.56$, $p=0.213$, and interactions, $F(1, 153)=0.70$, $p=0.791$, significance was not achieved suggesting that the changes in the dependent variable may be a result of the sample rather than the intervention. ANOVA 2 followed a similar pattern with only a main effect of time, $F(1, 181)=23.53$, $p<0.001$, $\eta_p^2=0.115$. Consequently, no further testing was performed on ANOVA 1 and 2. ANOVA 3 also had a significant main effect of time, $F(1, 133)=10.99$, $p<0.001$, $\eta_p^2=0.076$ and condition, $F(1, 133)=4.83$, $p=0.030$, $\eta_p^2=0.035$. However, more importantly, there was no significant interaction, $F(1, 133)=0.74$, $p=0.390$, indicating that the changes in pre-test and post-test scores were not influenced by the intervention. Overall, the results suggest that there was no significant effect of the intervention on mental arithmetic scores for either quintile.

7.7 AfE (InCAS) assessments -Reading and SIMD

7.7.1 Word recognition subtest

To ensure that there were no significant differences between the two groups at the pre-intervention stage, independent *t*-tests were performed on the pre-test scores of learners from the most and least deprived quintiles. For SIMD 1 learners, no significant differences were found, $t(153)=-1.28$, $p=0.202$. Examination of the mean differences between the groups revealed that the control group was approximately one year and one month behind their age-appropriate level ($M=-1.09$, $SD=1.4$) and the intervention group had approximately an eight-month deficit ($M=-0.70$, $SD=1.7$). Inspection of the mean scores for SIMD 5 learners in the control group revealed they were one and a half months behind their expected level, ($M=-0.15$, $SD=2$) and the intervention classrooms were three-months ahead of their age equivalent level, ($M=0.28$, $SD=1.7$). Once again there were no significant differences between the groups before starting the intervention, $t(181)=-1.59$, $p=0.114$ and a poverty-associated attainment gap were observed.

	Within-Subject			Between-Subject			Interactions		
α=0.05	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
ANOVA 1 (1,153)	39.72	0.001	0.206	4.21	0.042	.027	4.47	0.036	0.028
ANOVA 2 (1,181)	37.73	0.001	0.172	6.45	0.012	.034	6.45	0.012	0.034
ANOVA 3 (1,133)	22.31	0.001	0.144	2.25	0.136	.017	4.67	0.032	0.034

Table 52: Results from the two-way mixed ANOVAs conducted on the word recognition subtest.

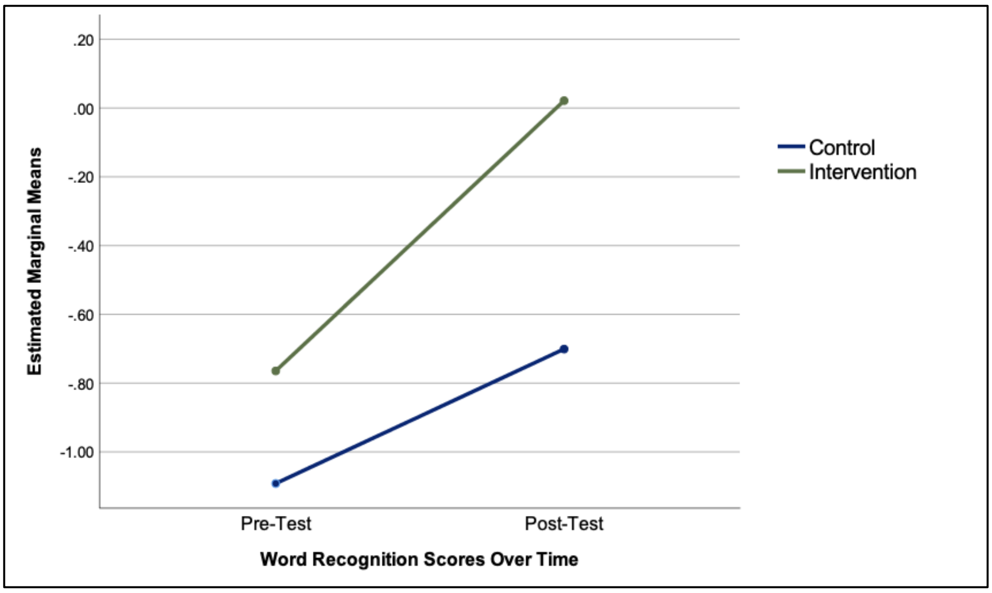
With reference to pre-test and post-test scores, results from ANOVA 1, presented in Table 52 revealed a significant main effect. The effect size for the difference was large. When the same analysis was conducted on the follow-up tests it revealed that both the control, $t(66)=-2.76$, $p=0.008$ and intervention groups, $t(87)=-6.70$, $p<0.001$ demonstrated progress from baseline. Inspection of the means indicates that the change was greater in the intervention group (MD=-0.786, SD=1.1) compared to the control (MD=-0.406, SD=1.2). In addition, a significant interaction was revealed between time and condition. As Figure 66A illustrates, the factor driving the interaction is the intervention post-test score increased to a greater degree than the control. Analysis of the differences between both groups at follow-up show that there was a significant difference between the two groups at the end of the intervention, $t(153)=-2.51$, $p=0.012$, suggesting that the change in post-test scores was affected by the classroom the learner attended.

The ANOVA was repeated, controlling for the acoustic environment. Interestingly, this revealed a significant interaction in the fair classrooms, $F(1, 56)=4.08$, $p=0.044$,

η_p^2=0.068, and a non-significant difference in the good classrooms, $F(1, 95)=1.75$, $p=0.189$. As Figure 66B illustrates, it is once again the post-test scores in the intervention classroom driving the interaction. In contrast to the results in numeracy and developed ability where the dynamic soundfield was primarily effective in good speech weighted C_{50} classrooms, the word recognition subtest only demonstrated a significant improvement in the fair C_{50} classrooms. Thus, learners from SIMD 1 are more likely to achieve improvements in word recognition scores in classrooms fitted with dynamic soundfield that have suboptimal acoustic conditions for speech.

Classrooms that are categorised as fair for speech will not have as many early reflections as the good classrooms (Bradley et al., 2003). The C_{50} mean for the intervention classrooms was 2.02dB (SD=0.82dB). One explanation for the improvement in scores is that the SIMD 1 learners in the fair speech weighted C_{50} classrooms obtained a beneficial effect of the adaptive gain provided by the dynamic soundfield, which mitigated the effects of the poor acoustics. The fair intervention classrooms in this research had a mean RT_{60} of 0.97s (SD=0.25s), with mean noise levels during literacy lessons of 63.67L_{Aeq} (SD=2.18 L_{Aeq}). The experimental testing by Dance et al. (2018) showed that the dynamic soundfield in similar levels of noise and RT_{60} provided an equivalent noise reduction level of 6dB. This would suggest that the dynamic soundfield enhanced speech intelligibility and overcame the limitations of rooms that were suboptimal for early reflections. Overall, the results suggest that dynamic soundfield had a positive effect on the word recognition scores for learners from the most deprived quintile.

A)



B)

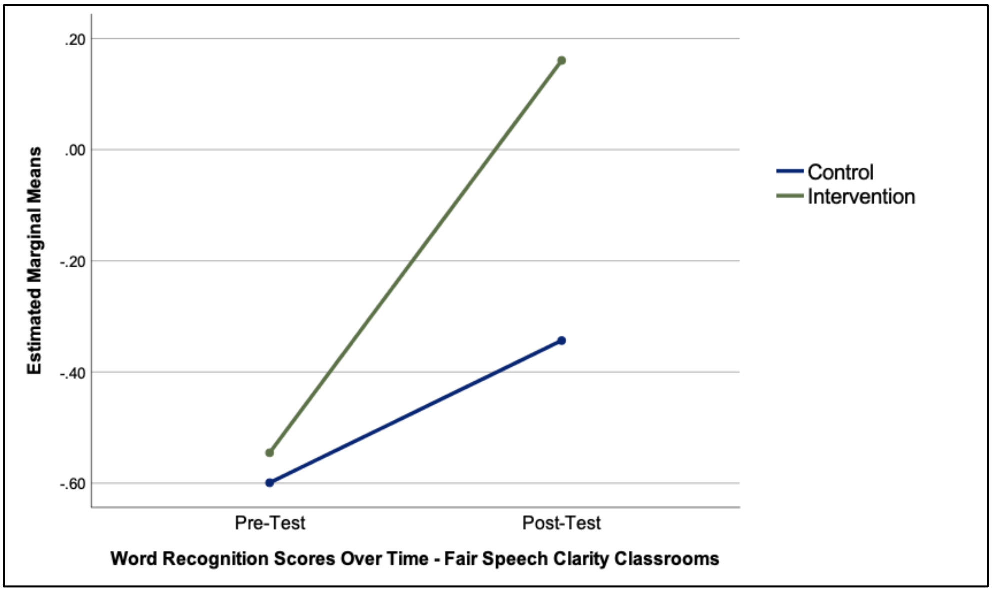
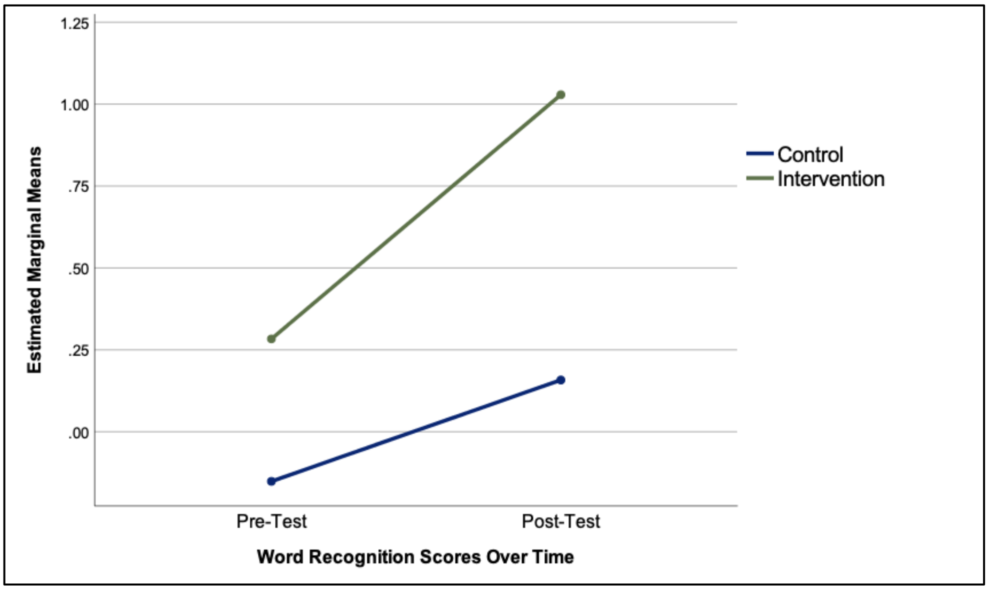


Figure 66: A) Interactional line chart showing the change in word recognition scores from the pre-test and post-test assessments for SIMD 1 learners. B) ANOVA repeated in classrooms fair for speech clarity (n=5).

Consistent with the analysis of ANOVA 1, ANOVA 2 also identified a significant time by condition interaction. As Figure 67A illustrates the post-test scores in the intervention group (green line) increased at a greater rate than the control group (blue line). Inspection of the means confirms that the intervention group improved the most, making significant gains from baseline, $t(103)=-9.03$, $p<0.001$. In comparison the control classrooms made non-significant, modest progress, $t(78)=1.88$, $p.=0.064$. Analysis of the means indicates that the intervention group (MD=-0.745, SD=1) was driving the interaction with larger gains than the control classrooms, (MD=-0.309, SD=1.5). Follow-up independent *t*-tests confirmed that there was a significant difference between the control and intervention groups at the end of the experimental phase, $t(181)=-3.24$, $p<0.001$.

The ANOVA 2 test was repeated, controlling for the two acoustic environments. Only classrooms good for speech clarity demonstrated a significant main effect of time, $F(1, 134)=22.21$, $p<0.001$, $\eta_p^2=0.142$, condition, $F(1, 134)=6.87$, $p=0.010$, $\eta_p^2=0.049$, and interaction, $F(1, 134)=4.73$, $p=0.031$, $\eta_p^2=0.034$. As Figure 67B illustrates, collectively the significant change in post-test scores in the intervention classrooms with good C_{50} values produced the significant interaction. Interestingly, in contrast to ANOVA 1, the results suggest that the dynamic soundfield intervention is most effective for word recognition scores in classrooms with good speech clarity properties.

A)



B)

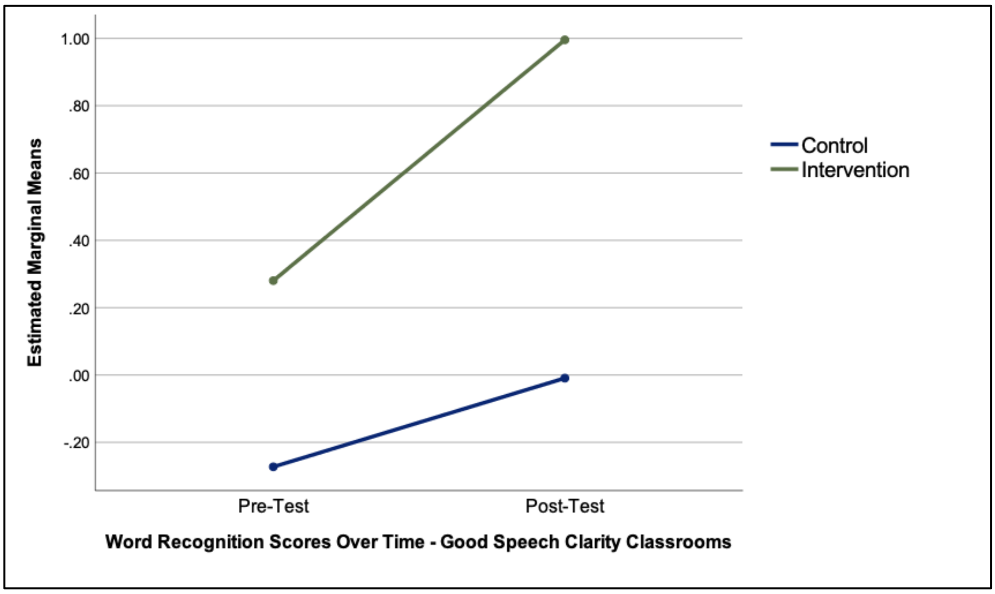


Figure 67: A) Interactional line chart showing the change in word recognition scores from the pre-test and post-test assessments for SIMD 5 leaners. B) ANOVA repeated in classrooms good for speech clarity (n=18).

Consistent with the results from the first two ANOVAs, ANOVA 3 also revealed a significant main effect of time and an interaction of time x condition, see Table 52. Figure 68 graphically illustrates the interaction. To establish which of the groups was driving the interaction follow-up tests were performed. Results revealed that SIMD 1 learners in the intervention classrooms with good speech clarity made greater (MD=-0.832, SD=1.3) and significant, $t(55)=-4.93$, $p<0.001$ progress on post-test scores compared to SIMD 5 learners in the control. An independent *t*-test also found that at baseline there was a significant difference between the groups, $t(133)=-2.19$, $p=0.030$ and at the end of the intervention this was non-significant, $t(181)=-0.628$, $p=0.531$, suggesting that the gap between the two groups had closed. Although maturation may be a factor as SIMD 1 learners were approximately six-months behind SIMD 5 learners in the control classrooms at the start of the study, it does not fully explain the difference. As discussed, the results from ANOVA 1 confirmed that SIMD 1 learners in the intervention classrooms did significantly better than their counterparts in the control, indicating a differential effect for the dynamic soundfield classes. Overall the results suggest that the dynamic soundfield intervention contributed to closing the gap in word recognition scores between SIMD 1 learners in the intervention classroom and SIMD 5 learners in the control.

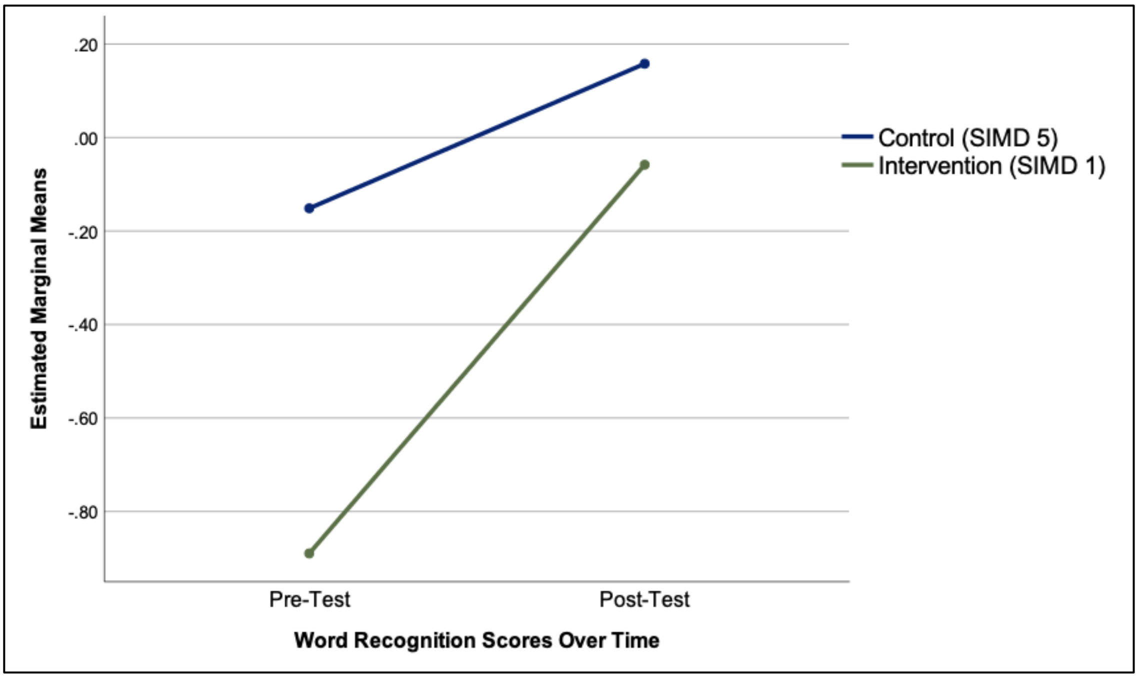


Figure 68: Interactional line chart showing the change in word recognition scores from the pre-test and post-test assessments.

7.7.2 Word Decoding subtest

There were no significant differences between the SIMD 1 learners in the intervention classes and the comparison group, $t(153)=-1.65$, $p=0.101$. A similar outcome was observed in the least deprived quintile, $t(182)=-0.753$, $p=0.452$.

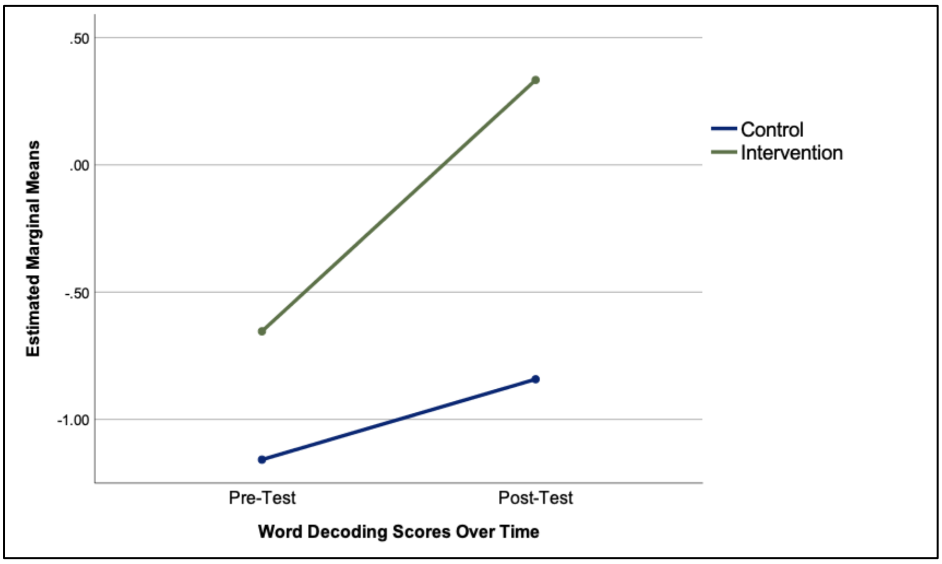
	Within-Subject			**Between-Subject**			**Interactions**		
α=0.05	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
ANOVA 1 (1,153)	26.03	0.001	0.145	8.05	0.005	0.050	6.91	0.009	0.043
ANOVA 2 (1,181)	64.59	0.001	0.263	2.60	0.109	0.014	2.63	0.107	0.014
ANOVA 3 (1,133)	40.74	0.001	0.235	4.39	0.038	0.032	0.287	0.593	0.002

Table 53: Results from the two-way mixed ANOVAs conducted on the word decoding subtest.

As Table 53 illustrates, a highly significant pre-test and post-test main effect were identified for ANOVA 1. In the control group, significant improvement was not achieved between the pre-trial stage and follow-up, $t(66)=-1.70$, p=0.094. In contrast, the learners exposed to the intervention demonstrated a significant improvement in scores at the follow-up assessment compared to the baseline, $t(87)=-5.73$, p<0.001. Inspection of the means confirmed that progress in the control group (MD=-0.316, SD=1.5) was lower than in the intervention classes (MD=-0.988, SD=1.6). There was also a significant main effect of condition, with post-test scores showing that there was a significant difference between the two groups by the end of the academic year, $t(153)=-3.47$, p<0.001. Consistent with this analysis is a significant condition x time interaction (Figure 69A) which shows that the intervention classrooms (green line) made greater progress over the year and this resulted in a significant difference between the groups at the post-intervention stage. Both these factors drove the interaction.

The tests were repeated controlling for the different speech-weighted C_{50} classroom environments and a significant main effect of time, $F(1, 95)=11.01$, $p<0.001$, $\eta_p^2=0.104$ condition, $F(1, 95)=11.92$, $p<0.001$, $\eta_p^2=0.111$ and interaction $F(1, 95)=7.81$, $p=0.006$, $\eta_p^2=0.076$ were identified in classrooms categorised as good for speech. Fair classrooms showed no significant interaction effect. As Figure 69B illustrates one factor driving the interaction was the significant increase in the intervention subtest scores, which are attributable to the good classroom environment and a modest non-significant increase in the control classroom scores. Interestingly, this suggests that the control classroom with fair acoustics performed better than the good classroom condition. Furthermore, the similarity between the trajectory of the green line (intervention group) in Figure 69B and 69A indicates that the significant increase in scores was primarily attributable to the good C_{50} classrooms. The mean C_{50} values in the intervention classrooms was 4.55dB (SD=0.96dB). Collectively, this indicates that these two factors produced the interaction effect.

A)



B)

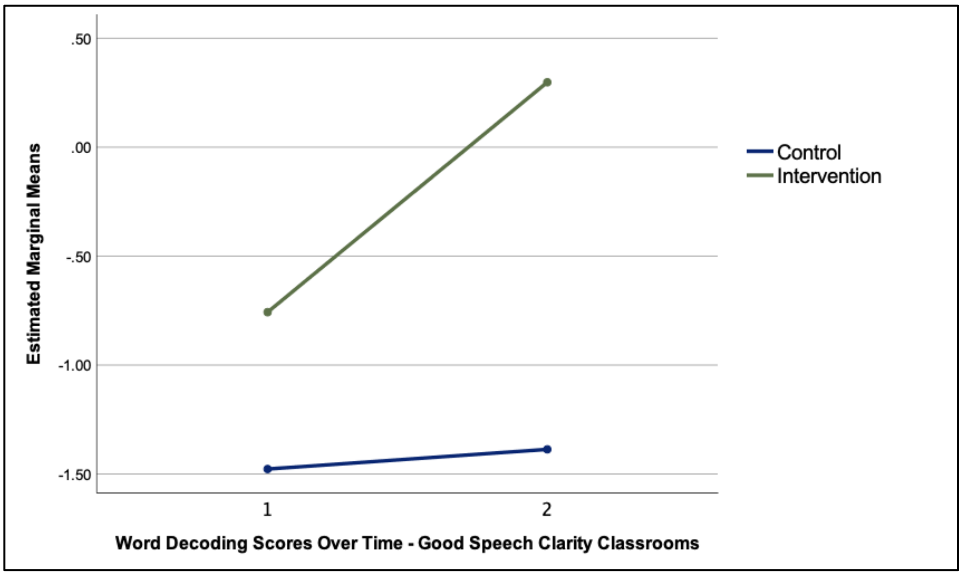


Figure 69: A) Interactional line chart showing the change in word decoding scores from the pre-test and post-test assessments. B) Mixed ANOVA repeated in classrooms good speech clarity (n=18).

As Table 53 illustrates only a significant main effect of time was observed in ANOVA 2. The large effect size is primarily due to both the control, *t*(78)=-4.51, p<0.001 and intervention groups, *t*(103)=-7.06, p<0.001 makings significant progress from baseline. The intervention group made greater progress (MD=1.34, SD=1.9) compared to the control, (MD=0.892, SD=1.8). Both condition and interaction did not achieve significance and so the differences observed are more likely a factor of the sample and cannot be attributed to the intervention. No further testing was performed. Consistent with these results, ANOVA 3 also revealed that there was a significant effect of time, $F(1, 133)=40.75$, $p<0.001$, $\eta_p^2=0.235$. Both SIMD 1 learners in the intervention and SIMD 5 learners in the control demonstrated significant progress from baseline. It is noteworthy that the interaction between condition and time was not significant. Once again, no further testing was performed. Overall these findings would suggest that the improvements in word decoding scores were only observed by SIMD 1 learners and this was mainly in classrooms that had good speech clarity.

7.7.3 Reading Comprehension subtest

No significant differences between the control and intervention groups were discerned at baseline for SIMD 1 learners, $t(153)=-1.76$, $p=0.080$. A similar outcome was observed for SIMD 5 learners, $t(181)=-1.35$, $p=0.179$. Furthermore, all three ANOVAs did not achieve significance for the interaction between time and condition. ANOVA 1 revealed a significant main effect of time, $F(1, 153)=34.45$, $p<0.001$, $\eta_p^2=0.184$ which follow up tests revealed was driven by both the control, $t(66)=2.51$, $p=0.015$ and intervention groups, $t(87)=-6.16$, $p<0.001$ making significant progress from time point 1 to time point 2. ANOVA 2 followed a similar pattern, $F(1, 181)=48.30$, $p<0.001$, $\eta_p^2=0.211$ with both the control, $t(78)=-3.01$, $p=0.004$ and intervention group, $t(103)=-7.50$, $p<0.001$ achieving significance. Consistent with these results, ANOVA 3 also revealed a significant main effect of time, $F(1, 133)=26.20$, $p<0.001$, $\eta_p^2=0.165$. Although all three ANOVAs identified a main effect of condition, there was no significant interaction and so the outcome cannot be attributed to the intervention. No further testing was performed.

7.7.4 Spelling subtest

The groups did not significantly differ in terms of the baseline assessment. There was a non-significant difference between SIMD 1 learners in the control and intervention classrooms, $t(153)=-1.63$, $p=0.106$. This was replicated for the SIMD 5 learners, $t(181)=-1.29$, $p=0.199$. Spelling had a similar outcome to that observed in the Reading Comprehension subtest: ANOVA 1 had a main effect of time only, $F(1, 153)=54.41$, $p<0.001$, $\eta_p^2=0.262$. Both condition, $F(1, 153)=2.81$, $p=0.093$ and interaction, $F(1, 153)=0.04$, $p=0.893$ were non-significant. ANOVA 2 had a similar outcome: time, $F(1, 181)=95.95$, $p<0.001$, $\eta_p^2=0.346$, condition $F(1, 181)=3.62$, $p=0.059$ and interaction $F(1, 181)=2.83$, $p=0.094$. ANOVA 3 also only achieved significance on the main effect of time, $F(1, 133)=48.67$, $p<0.001$, $\eta_p^2=0.268$. The results from all three ANOVAs are consistent with a longitudinal study into education attainment, in which progress is made between the start and end of the term. Both condition, $F(1, 133)=3.61$, $p=0.060$ and interaction, $F(1, 133)=0.70$, $p=0.444$ did not achieve significance. No further testing was performed.

7.7.5 Reading module

Standardised reading scores at baseline confirms that there was no significant difference between the groups of SIMD 1 learners, $t(153)=-1.92$, $p=0.060$ and SIMD 5 learners, $t(181)=-1.25$, $p=0.212$. Inspection of the means showed that the gap identified between SIMD 1 and 5 learners at the start and end of primary 1, was still present at the start of primary 3. SIMD 1 learners in the control ($M=89.91$, $SD=12$) and intervention classrooms ($M=94.12$, $SD=14$) were behind their SIMD 5 counterparts in both the control ($M=100.81$, $SD=17$) and intervention classes ($M=103.88$, $SD=16$).

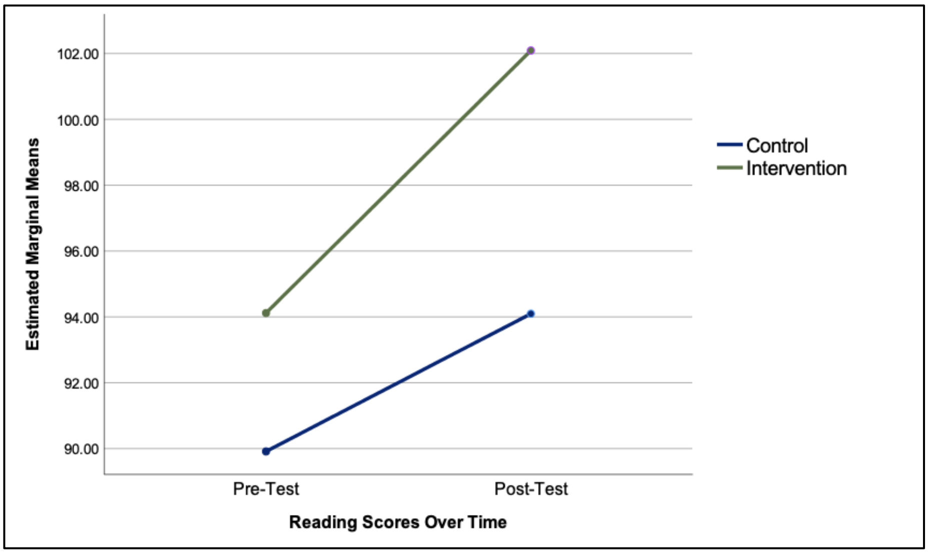
	Within-Subject			Between-Subject			Interactions		
α=0.05	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
ANOVA 1 (1,153)	59.90	0.001	0.145	7.45	0.006	0.048	5.83	0.017	0.037
ANOVA 2 (1,181)	78.03	0.001	0.301	5.04	0.026	0.027	8.34	0.004	0.044
ANOVA 3 (1,133)	37.48	0.001	0.220	3.78	0.054	0.028	3.02	0.085	0.022

Table 54: Results from the two-way mixed ANOVAs conducted on the standardised reading scores module.

To answer the research question of whether learners from SIMD 1 in the intervention classrooms perform better in reading than their peers in the control, ANOVA 1 was conducted on the standardised reading scores. As Table 54 illustrates there was a significant interaction between time and condition, indicating that one of the group's scores had increased to a greater degree than the other. Follow up tests on the effects confirm that both the control, $t(66)=-3.50$, $p<0.001$ and intervention groups, $t(87)=-7.79$ $p<0.001$ made significant progress during the study. The progress was greater in the intervention classrooms (MD=-7.97, SD=9.6) compared to the control (MD=-4.18, SD=9.7). As Figure 64A illustrates the differential change in pre-test and post-test scores in the groups was driving the interaction. Independent *t*-tests confirmed this, showing that there was a significant difference between the groups at the follow-up assessment, $t(153)=-3.25$, $p<0.001$ compared to a non-significant difference at baseline, $t(153)=-1.92$, $p=0.060$.

Controlling for the acoustic conditions found a significant interaction in the good classrooms for speech, $F(1, 95)=4.39$, $p=0.039$, $\eta_p^2=0.165$ and a non-significant interaction in the fair classroom environments, F(1, 56)=0.212, p=0.647. As Figure 64B reveals, once again it is the post-test scores in the intervention classroom driving the interaction. Overall, the results suggest that learners from the most deprived quintile gained a significant benefit of dynamic soundfield for reading in classrooms that were categorised as having good speech clarity.

A)



B)

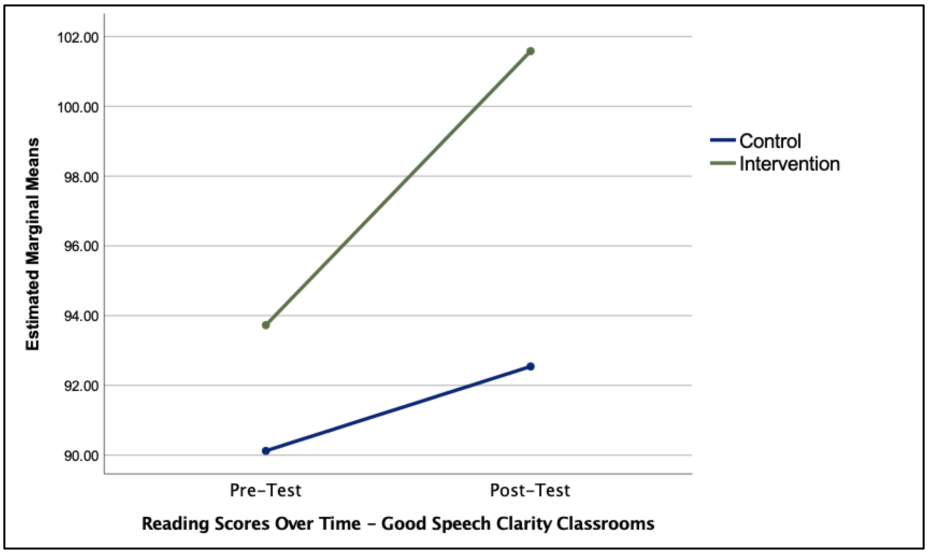
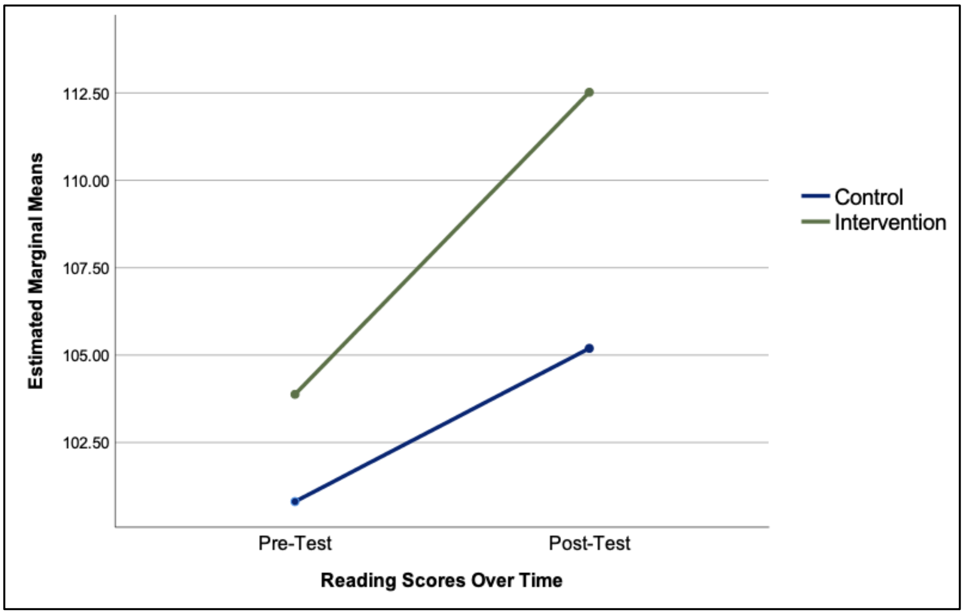


Figure 70: A) Interactional line chart showing the change in reading standardised scores for SIMD 1 learners. B) Mixed ANOVA repeated in good classrooms for speech clarity (n=18).

As Table 54 reveals, ANOVA 2 also achieved a significant main effect of time, condition, and more importantly an interaction. Follow up tests confirmed that both the control, $t(78)=-3.29$, $p=0.002$ and intervention classrooms, $t(103)=-10.93$, $p<0.001$ made significant progress between the pre-test and post-test time points. Once again, the progress was greater in the intervention classroom (MD=8.64, SD=9) with learners exposed to dynamic soundfield making approximately double the improvement compared to the control (MD=4.38, SD=11). The results are consistent with ANOVA 1 where a similar amount of progress was recorded in both the control and intervention classrooms. Independent *t*-tests were run to decompose the interaction further and this revealed that there was a significant difference between the groups at the end of the experimental phase, $t(181)=-3.04$, $p=0.003$. This is graphically illustrated in Figure 71A.

To determine the influence of classroom acoustics on the efficacy of the dynamic soundfield, ANOVA 2 was repeated comparing the gains made in both the good and excellent listening environments. The results in the good classrooms for speech also presented a similar significant main effect of time, $F(1,134)=48.76$, $p<0.001$, $\eta_p^2=0.267$ condition, $F(1,134)=5.48$, $p=0.021$, $\eta_p^2=0.039$ and interaction, $F(1,134)=4.95$, $p=0.028$, $\eta_p^2=0.036$. The similarities of the trajectory in the intervention classrooms between time points 1 and 2 graphically illustrated in Figure 71A and Figure 71B suggests that it was primarily a good classroom environment for speech driving the change in scores. There were no significant interaction effects in the excellent classrooms, $F(1,44)=3.91$, $p=0.064$. Once again, these findings would indicate that although progress was observed in the intervention classrooms throughout the study, the improvements were primarily in classrooms that were categorised as being good for speech clarity.

A)



B)

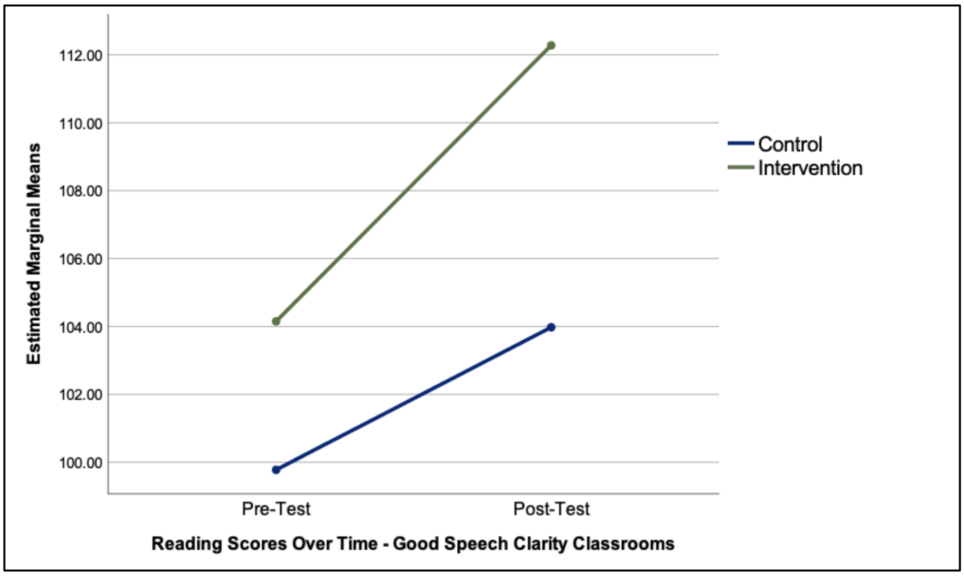


Figure 71: A) Interactional line chart showing the change in reading scores from the pre-test and post-test assessments for SIMD 5 learners. B) Mixed ANOVA repeated in good classrooms for speech clarity (n=18).

ANOVA 3 revealed a significant main effect of time, which follow up tests revealed was driven by both the control, $t(78)=-3.29$, p=0.002 and intervention groups, $t(55)=4.23$, p<0.001 making significant progress from time point 1 to time point 2. Table 54 reveals both the condition and interaction effect was non-significant. This suggests that dynamic soundfield did not significantly affect the difference between SIMD 1 learners in the intervention classroom and SIMD 5 learners in the control. No further testing was performed.

7.7.6 Summary – Reading module and subtests

The results of this study suggest that young people from the most and least deprived quintiles benefit from dynamic soundfield systems in the classroom. In both the Word Recognition subtest and standardised Reading scores, SIMD 1 and SIMD 5 learners demonstrated a beneficial effect of the dynamic soundfield system. The beneficial effects of good acoustics are not enough to mitigate the detrimental effects that young people experience when listening to speech in noise. It may be that the distinct amplified voice of the teacher, provided by the dynamic soundfield allows the listener easier access to the target signal of interest and so reduces listening effort.

Only SIMD 1 learners in the intervention classrooms demonstrate a significant improvement in the Word Decoding subtest. The findings coincide with previous research which indicates that young people go through a period of auditory maturation, with the ability to discriminate consonants in noise not fully mature until the teenage years (Johnson, 2000). As discussed in Chapter 2, this appears to disproportionately disadvantage young people from areas of social deprivation and those with additional support needs (Bhang et al., 2018).

Only SIMD 1 learners demonstrated a significant improvement in the Word Recognition subtest scores in classrooms categorised as fair for speech clarity. All other significant results showed improvement only in classrooms with good C_{50} properties. In the Word Recognition subtest, the results showed that a significant gap between SIMD 1 and 5 learners at baseline was non-significant at the end of the intervention, indicating that the poverty associated attainment gap between the two groups was significantly reduced.

7.8 Discussion of attainment and deprivation results

Although several studies (Arnold and Canning, 1999, Rosenberg et al., 1999, Darai, 2000, Massie et al., 2004, Massie and Dillon, 2006b, Heeney, 2007) have suggested that soundfield amplification is effective at improving educational outcomes for school-aged learners, the current study is unique as no study has examined the effects on young people from the most and least deprived backgrounds, based on the individual's home postcode. Many of the studies into the effects of noise, control for socioeconomic status. Charlotte and Katarina (2018) completed a systematic review of thirty-four papers examining the effects of environmental noise on cognition in young people and found that socioeconomic factors were generally adjusted for as a potential confounding variable. One of the catalysts for the current study was to begin to address this gap in the literature, with the primary aim of tentatively examining whether dynamic soundfield is an effective intervention to improve outcomes for learners from SIMD 1, and so reduce the attainment gap.

The findings from the current study suggest that in the Non-Verbal Ability subtest there was a significant reduction in the gap between SIMD 1 learners exposed to dynamic soundfield and SIMD 5 learners in the control. The results from the baseline Non-Verbal Ability subtest concurs with previous research by Bradshaw (2011) which found that young people at the age of five from low-income families displayed a ten-month delay in non-verbal reasoning compared with the highest income group. The gap was larger, thirteen-months, for families with no qualifications compared to a degree educated parent. In this research, the post-test analysis demonstrated a significant improvement in scores in the classrooms exposed to dynamic soundfield.

One possible explanation is that vulnerable groups, such as learners living in areas of social deprivation are disproportionately affected by noise. Studies have shown that noise can have a detrimental effect on tasks that are commonly associated with cognitive function (Cohen et al., 1980) and it may be that the amplified classrooms provide the listener with easier access to the target signal of interest and so reduce listening effort. Although there is a paucity of literature on which this hypothesis can be compared, the results from this study broadly mirror findings from Bhang et al. (2018) who reported that noise compromised attention and cognitive function, especially affected were young people with low academic performance.

Findings also reveal that there was a significant reduction in the gap between SIMD 1 and 5 learners in the Word Recognition subtest. At pre-test, SIMD 1 learners in both conditions presented with a significant delay in word recognition capacity compared to SIMD 5 learners. Two trends were observed. Firstly, both SIMD 1 and 5 learners in the intervention classrooms demonstrated a significant improvement in word recognition ability, with both making approximately eight-months progress from baseline. It is noteworthy that both groups made comparable progress, as it was anticipated that learners with the largest delay in word recognition ability would have demonstrated the greatest improvement. This suggests there is a beneficial effect of dynamic soundfield on word recognition performance regardless of your SIMD categorisation. Secondly, SIMD 1 learners exposed to dynamic soundfield made significant progress in word recognition compared to SIMD 5 learners in the control classrooms. Together these findings align with previous studies which indicate that young people’s capacity to listen in noise is age-related (Klatte et al., 2010a, Wightman et al., 2010). The significant interaction identified in the two-way ANOVA also indicates that the improvement in post-test scores was attributed to learners attending the adaptive amplified classrooms.

Interestingly, the only subtest where the dynamic soundfield system was effective in classrooms with fair C_{50} properties was word recognition for SIMD 1 learners. One possible explanation may be related to the noise levels and reverberation times at which the dynamic soundfield system becomes beneficial for speech intelligibility. The mean RT_{60} in the fair classrooms for speech was 0.967s (SD=0.25s) In the experimental research by Dance et al. (2018), the dynamic soundfield system was only beneficial in classrooms with longer reverberation times when interfering noise levels reached a certain point. In rooms with RT_{60} ≥1.0s, the dynamic soundfield system provided an ENR of 6dB when noise levels reached 65dB(A). In Chapter 5 the results from the noise surveys in classrooms revealed that literacy recorded the lowest noise levels compared to IDL and numeracy. Overall noise levels were 62.79 L_{Aeq} in the control and 63.67 L_{Aeq} in the intervention classrooms. The most commonly observed learning activities (A3-A5) had average noise levels of 64.2 in the control and 63.8 L_{Aeq} in the intervention. The positive results in the word recognition subtest suggest that even at lower noise levels in rooms with fair speech clarity the listener gains a benefit from amplification compared to the equivalent control. This again may be a result of the dynamic soundfield system spatially separating the target signal of interest from the masker and so providing access to the cues that help the listener filter speech in noise.

The findings on SIMD 1 learners and word decoding are also interesting, showing that in amplified classrooms with good C_{50} values there was a significant improvement in scores compared to the control. Both the control and intervention groups were delayed in their word decoding capacity at baseline. This ties in with research by Nittrouer and Burton (2005) who found learners from lower socioeconomic backgrounds have delays in the perceptual strategies required to extract phonetic structure from the acoustic signal of interest. The development of these perceptual strategies is nurtured through extensive experience, with delays in early language experience compromising speech perception and phonological processing. It might be expected that young SIMD 1 learners whose performance is significantly poorer would gain a benefit from amplified classrooms as the target signal of interest would be distinct from the multiple competing voices in the room. The results from the current study suggest that this may be the case, as there was a significant improvement in the word decoding capacity amongst SIMD 1 learners exposed to dynamic soundfield. The current study, therefore, contributes to the limited knowledge base on the effects of amplification on the perceptual strategies that support word decoding.

In addition to word recognition and word decoding, results from the current study indicate that there was a significant improvement in the picture vocabulary tariff for SIMD 1 learners exposed to the intervention. At baseline, both the control and intervention groups were behind their age equivalent level which is consistent with the findings from Bradshaw (2011) who identified a significant difference in vocabulary ability levels between learners whose parents had higher and lower education qualifications and learners from the lowest and highest income groups. At the follow-up assessment, the classrooms exposed to dynamic soundfield demonstrated a significant improvement in mean scores, in contrast, the control classrooms showed no significant change. The disordinal interaction indicates that the improved change in scores was significantly associated with attending the intervention classrooms. One explanation for this positive change is that it is easier to identify the target signal of interest in noise when there is an improved SNR, resulting in the listener receiving a beneficial target signal of interest to the masker ratio. Another possible explanation is that it is easier for the listener to separate the amplified voice of the teacher from the interfering noise and this release from masking makes it easier for young people with significant language delays to access the spoken word. Although there are limited studies involving young people from areas of social deprivation to compare these hypotheses, the findings are consistent with previous research by Boothroyd (1997) which shows that the auditory speech perception system is strongly associated with linguistic knowledge and experience. Without robust auditory templates, the listener may be disproportionately disadvantaged when listening in noise.

Both SIMD 1 and 5 learners in the intervention classrooms demonstrated a significant improvement in their overall reading scores. As discussed in the previous chapter, the simple view of reading is decomposed into two connected components; word decoding and language comprehension. Findings from the current study show there was no significant improvements in reading comprehension for both quintiles, this suggests that the improved performance in reading was primarily a result of improved word decoding/recognition capacity. For SIMD 1 learners, the significant improvement in picture vocabulary scores may also be a contributing factor.

One of the skills commonly associated with successful reading development is the capacity to separate spoken language into sound segments. This includes large sound units, words, and syllables to intermediate and smaller units such as onset, codas, rimes and phonemes (Melby-Lervåg et al., 2012). A body of research (Nittrouer and Boothroyd, 1990, Johnson, 2000, Nittrouer, 2002) has shown that the ability to identify and discriminate the different parts of speech in noise is age-related. The findings from this study align with previous research as both SIMD 1 and 5 in the intervention classrooms demonstrated significant improvements compared to the control. This suggests that improvements in word recognition were not specific to socioeconomic factors. Only SIMD 1 learners demonstrated a significant improvement in word decoding. The results from the primary 1 PIPS demonstrated that SIMD 1 learners were delayed in their phonological processing skills compared to SIMD 5 learners. This suggests that improvements in word decoding were specific to social deprivation factors. The changes in the pre-test/post-test scores observed in the reading module and subtests in the intervention classrooms suggests a dynamic soundfield advantage in this group compared to the control. One possibility might be that the amplified voice of the teacher provides a higher target to masker ratio and so mitigate the effects of informational masking (Wightman and Kistler, 2005).

Extracting meaning from the acoustic signal is an important aspect of reading. Although there was no differential socioeconomic effect for the reading module, it is noteworthy that SIMD 1 learners exposed to dynamic soundfield performed better than their counterparts in the control. As presented earlier in this thesis, there was a poverty-associated attainment gap in phonological processing both during the first year of school and at the start of this study. SIMD 1 learners exposed to dynamic soundfield demonstrated a significant improvement in the Word Decoding and Word Recognition subtests, which are two of the four subtests used to calculate the overall reading module score. It is well-documented that there is a national poverty associated reading gap (Jerrim, 2013, Sosu and Ellis, 2014, Scottish Government, 2016c) and these findings suggest that dynamic soundfield contributes towards improved outcomes in reading.

The current study found that only learners from SIMD 5 showed a significant improvement in the general mathematics subtests of data handling and measure, shape and space. In both subtests, the significant improvements were only in the intervention classrooms with excellent C_{50} values. The mean C_{50} value of the excellent intervention classrooms was 7.53dB, approximately 3dBs higher than in the good C_{50} classrooms. However, as discussed previously the small and unrepresentative sample size means the findings should be treated with caution. The reasons that SIMD 5 learners exposed to dynamic soundfield were the only quintile to demonstrate a significant improvement are unclear, although it is noteworthy that in primary 1, SIMD 5 learners made the least amount of progress during the year in the early mathematics module. One possible explanation for the dynamic soundfield advantage is that these subjects require a higher level of multitasking, which consumes a larger amount of cognitive resources and so increases listening effort. Dual-task paradigms are generally used to measure listening effort which involves the participants repeating words (primary task) in noise whilst simultaneously rehearsing a set of digits to recall (secondary task). Past research has demonstrated that higher performance in the language focused primary task is accompanied by a decrement in the numeracy related secondary task (Choi et al., 2008, Howard et al., 2010). This is consistent with the finding from Hällgren (2005) who observed that when the target signal of interest is compromised by noise then the listening situation becomes more cognitively demanding and relies more on working memory capacity and selective attention. One possible explanation for the improvement in the intervention group is that the reduced amount of listening effort afforded by a higher SNR generated by the dynamic soundfield in classrooms with excellent C_{50} properties. Combined, this may have provided additional cognitive resources for tasks that involve evaluation and rapid information processing.

7.9 Key Findings

The primary aim of this chapter was to answer the primary research questions 2-5 of this study (see section 3.2.1).:

- To establish if the gap in educational attainment between Primary 3 learners from the most and least deprived quintiles reduced after being exposed to dynamic soundfield.
- To investigate if the changes in scores for each of the subtests and modules were significantly different between SIMD 1 learners exposed to dynamic soundfield and SIMD 1 learners in the control.
- To investigate if the changes in scores for each of the subtests and modules were significantly different between SIMD 5 learners exposed to dynamic soundfield and SIMD 5 learners in the control.
- To investigate if the efficacy of dynamic soundfield is moderated by the speech weighted C_{50} properties of the classrooms.

This research was unique as the test results were graded by a third-party blind to the intervention. The sample size was large, and access to the results achieved in the first year of school provided context. The key findings are:

- A poverty-associated attainment gap already exists at the start and end of primary 1. The gap is more pronounced in early reading than early mathematics.
- There were no significant differences between the SIMD 1 and 5 learners in the control and intervention groups in literacy and numeracy during the first year of school.
- Learners from SIMD 5 recorded higher test scores at baseline in literacy, phonological awareness, and mathematics than SIMD 1 learners achieved at the end of the first full year of schooling.
- The analysis revealed that there was a significant reduction in the attainment gap between SIMD 1 learners in the intervention classroom and SIMD 5 control for the Non-Verbal Ability and Word Recognition subtests.
- It was only SIMD 1 learners in the amplified classrooms categorised as good for speech that gained a benefit from Picture Vocabulary subtest and the Developed Ability module. Although the gap between SIMD 1 and 5 learners was reduced, this was not significant.
- SIMD 1 learners in the intervention classrooms with good C_{50} values demonstrated significant improvements in the Non-Verbal Ability subtest compared to the control.
- Both SIMD 1 and 5 learners in the intervention classrooms gained a significant benefit in the Word Recognition subtest. SIMD 1 learners only gained a benefit in the fair classrooms. In contrast, SIMD 5 learners only gained a beneficial effect in classrooms with good C_{50} values. The analysis revealed that there was a significant reduction in the attainment gap between SIMD 1 learners in the intervention classroom and SIMD 5 control for the Word Recognition subtest.

- SIMD 1 and 5 learners in the intervention classroom gained a significant benefit in the Reading module, which was primarily in a good acoustic environment.
- Only SIMD 1 learners exposed to dynamic soundfield in rooms with good C_{50} values achieved significance in the Word Decoding subtest.
- Only SIMD 5 leaners exposed to dynamic soundfield in classrooms with excellent C_{50} values gained a significant benefit in the Data Handling and Measure, Shape and Space subtests.
- Mental Arithmetic showed no significant effect for the intervention in either quintile.
- Spelling and reading comprehension showed no significant differential effects of the intervention.
- Overall, the results suggest a significant benefit of the dynamic soundfield for learners from the most deprived quintiles in curricular areas associated with non-verbal cognitive tasks and those mediated through spoken language (reading, vocabulary, and decoding). Although some of the beneficial effects of dynamic soundfield were dependent on having a good acoustic environment, interpreting these results to draw conclusions about the efficacy of soundfield in excellent acoustic conditions, have to be treated with caution due to the unrepresentative and small sample size.
- Overall, SIMD 5 learners gained a beneficial effect of dynamic soundfield in classrooms that had good or excellent acoustics for speech. The benefit of non-verbal tasks (data handling, measure, shape, and space) was in classrooms with excellent acoustics for speech. The language tasks were effective in good acoustics for speech. Once again caution needs to be applied when drawing conclusions on the efficacy of soundfield in excellent acoustic environments.

Chapter 8

Conclusion and Contribution

8.1 General discussion and comparison with other studies

This thesis was mainly concerned with establishing whether there were significant improvements in educational attainment for Primary 3 learners exposed to dynamic soundfield amplification in their mainstream classrooms. Thereafter, to explore the interrelations between acoustic conditions, social deprivation status and adaptive amplification of the teachers' voice on attainment and the reduction in the poverty-associated attainment gap in the curricular areas of developed ability, literacy, and numeracy. As recalled from Chapter 2, previous studies have attempted to measure the efficacy of soundfield on educational improvements and speech intelligibility but were compromised by a poor study design that did not incorporate standardised assessments that were marked by an external body, blind to the intervention. There was also a paucity of research into the effects of an adaptive soundfield system on learners from the most and least deprived quintiles. These shortcomings from previous studies were part of the motivation for the current research.

This research contributes to the growing body of research which indicates that young learners are more susceptible to the effects of background and underlying noise, especially in suboptimal acoustic conditions (Blandy and Lutman, 2005, Klatte et al., 2005, Shield and Dockrell, 2008, Klatte et al., 2013). Chapter 4 investigated the ambient noise and RT_{60} levels in the 25 research classrooms and found that in general, these were above the recommended guidelines. Several studies both nationally and internationally have reported similar findings (Hodgson, 1994, Shield and Dockrell, 2004, Yee Choi and McPherson, 2005, Shield et al., 2015). Chapter 4 also discussed that early reflections of speech sounds are important to achieve adequate speech intelligibility in a classroom and presented a user-friendly rating scale that can be used to measure the speech weighted C_{50} properties of a classroom.

Chapter 5 examined the relationship between noise levels in occupied classrooms and learning activities commonly observed in a classroom. The results revealed that noise levels were highest for learning activities that involve multiple talkers and movement, which are activities that are a core component of an active learning environment. These findings are consistent with previous research in both primary and secondary schools (Shield and Dockrell, 2004, Shield et al., 2015). The overall mean noise levels during lessons were approximately 64dB(A) L_{Aeq}.

The most commonly observed noise internal to the school but external to the classroom were from other classrooms or other young people moving around the school. Learners from the most deprived quintile in the control classrooms were disproportionately affected by this type of noise. Chapter 4 showed that young people spend approximately 10 per cent of the school week not in the classroom. Furthermore, SIMD 1 learners in both the control and intervention classrooms were more aware of the noise generated by people retrieving things from bags and the scrapping of chairs. Combined these findings are important from a policy perspective since they contribute to the understanding of the effects that different teaching methods and timetabling have on noise levels in schools.

Chapter 5 demonstrated that the prediction of hearing the teacher easily in both low and high background noise levels was improved by access to dynamic soundfield amplification. The questionnaire results suggest that even when the majority of classrooms provided good C_{50} values and underlying noise levels were low, young learners identified they could hear the teacher easier in adaptive amplified classrooms. The scenario with the biggest effect size was when the class teacher was standing at the front of the room, noise levels were low, but no visual cues were present. In this situation, the direct sound is particularly weak, as the talker is turned away from the learner. Although early reflections should enhance speech intelligibility in such a scenario, the findings from the current study indicate these are not sufficient to make it easy for the listener to access speech. This suggests that young people require a combination of dynamic soundfield, positive C_{50} values and visual cues to supplement listening. This partially supports the hypothesis by Cherry (1953) that visual cues and speech characteristics provide a release from masking. It is unclear if the dynamic soundfield advantage is because of an improved SNR, the distinct characteristics of an amplified voice making it easier for the listener to identify spoken language or both.

Chapter 6 demonstrated there was a significant improvement in the Developed Ability subtests and standardised module scores for learners in the intervention classrooms. Learners exposed to dynamic soundfield improved their picture vocabulary knowledge, non-verbal skills and developed ability skills from baseline at a significantly higher level compared to the control. This was primarily observed in classrooms that were categorised as good for speech-weighted C_{50}. All the reading subtests apart from word decoding showed a significant benefit of the intervention in classrooms with fair and good speech clarity. The good speech-weighted C_{50mean} in the control classrooms were 4.85dB (SD=1.1dB) and 4.56dB (SD=0.97dB) in the intervention classrooms. This contrasts with a fair speech-weighted C_{50mean} of only 2.49dB (SD=0.52dB) in the control classrooms and 2.02dB (SD=0.82dB) in the intervention. Significant improvements in word decoding were only observed in classrooms with good C_{50} properties. Collectively, these findings suggest that young people gain an educational advantage by having the voice of the teacher amplified by a system providing adaptive gain. These findings are the first to show that the effectiveness of the dynamic soundfield system is conditioned by the acoustic environment using speech-weighted C_{50}.

Chapter 7 showed that SIMD 1 learners entered school behind their SIMD 5 counterparts in literacy, numeracy, and phonological processing. As discussed in Chapter 2 there is a paucity of research into the effects that dynamic soundfield has on educational attainment. Previous research has suggested that noise disproportionately affects learners that present with lower academic performance (Bhang et al., 2018). Although SIMD 1 learners demonstrated progress during the first year of school, SIMD 5 learners still recorded higher test scores at baseline than that achieved at the end of the first full year of school by SIMD 1 learners. These findings align well with past studies which reported a similar attainment gap in the pre-school years (Bradshaw, 2011, Sosu and Ellis, 2014). Unsurprisingly, the poverty-associated attainment gap was still present at the start of the current study.

Chapter 7 then showed that it was only learners from the most deprived quintile that gained a significant benefit for all the Developed Ability subtests. In keeping with the findings from Chapter 6, the dynamic soundfield was primarily effective in classrooms that were categorised as good for speech. This was the first study that appears to show that adaptive amplification in good C_{50} classrooms has a positive impact on the non-verbal ability and picture vocabulary knowledge for learners from the most deprived quintile. Together these findings support studies that suggest increased listening effort consumes greater cognitive resources that compromise educational performance. Conversely, higher SNR helps reduce listening effort which improves educational outcomes (Howard et al., 2010). The findings from this study extend the literature by demonstrating a differential effect for learners from the most and least deprived quintiles. Listening in noise is multidimensional with factors such as language, cognition and social deprivation influencing the process. These findings have implications for government and policymakers as the scores for Developed Ability are generally regarded as a good predictor for later academic achievement (Centre for Evaluating and Monitoring, 2014).

Chapter 7 also showed that the beneficial effect of dynamic soundfield for reading was associated with skills in word decoding/recognition capacity. Both SIMD 1 and 5 learners gained a benefit for the dynamic soundfield intervention in the Word Recognition subtest. It was only SIMD 1 learners that gained a benefit in the Word Decoding subtest in good C_{50} classrooms. It is interesting that SIMD 1 learners entered school with a phonological processing deficit compared to SIMD 5 learners and that this poverty-associated gap was still present at the start of primary 3. These results broadly support previous studies that have shown that young people from areas of social deprivation can have a delay in their phonological perceptual and processing capacity (Nittrouer and Burton, 2005, Bradshaw, 2011). SIMD 1 learners exposed to dynamic soundfield showed a significant improvement in word decoding in good C_{50} classrooms. This would suggest that the 4.56dB (SD=0.97dB) advantage provided by the early reflections in the good classrooms for speech clarity was insufficient to provide a release from masking. SIMD 1 learners, that had shown a significant deficit in their phonological processing skills since starting school, appear to require an effective increase in the SNR provided by dynamic soundfield and positive C_{50} values. These findings will be of interest to policymakers when making decisions on raising the attainment levels in reading. Decoding is a core component of reading and this research suggests that dynamic soundfield can help to contribute to improved performance for SIMD 1 learners.

Chapter 7 also illustrated that the poverty-associated attainment gap was significantly reduced in the non-verbal ability and word recognition subtests. Previous studies have shown that before starting school young people from areas of social deprivation have poorer vocabulary levels and problem-solving skills than their counterparts with higher socioeconomic classification (Bradshaw, 2011). The baseline assessments at the start of primary 3 showed that the poverty-associated attainment gap was still present after two years in school. At follow up, learners from SIMD 1 exposed to dynamic soundfield demonstrated significant improvements in non-verbal reasoning and word recognition. They made much greater gains than SIMD 5 learners in the control. Although maturation may have contributed to the outcomes, the differential results observed by SIMD 1 in the intervention classroom compared to the control demonstrated that dynamic soundfield was a significant contributory factor.

8.2 Strengths and limitations

A key strength of the study design was the large and representative sample size. As recalled from Chapter 4 the sample fairly represented both learners at the same primary 3 stage as the research cohort and all primary learners across Scotland. By establishing both internal and external validity it is possible to generalise the conclusions from this study. Allocation of the participants to the different control and intervention classrooms was by concealed randomisation and not through a classic random assignment, which is regarded as the gold standard. This would be one of the limitations of the current study. However, to achieve such random allocation in the current study design would have raised ethical and practical issues for the management in the schools. The management in a school generally decides the composition of a class grouping based on various factors such as age and characteristics of the learners. One of the advantages of random assignment is that any confounding variables are evenly distributed amongst the two conditions (Torgerson, 2008). As was evident in Chapter 4, there were no significant differences between the control and intervention groups in terms of participants' characteristics and acoustic environment.

A further strength was the small attrition rate achieved during this longitudinal study. Attrition can introduce bias and result in an unrepresentative sample size (Torgerson, 2008, Amico, 2009). As discussed in Chapter 4, the number of participants and teachers that dropped out of this research were below the five per cent level and so any bias would be minimal. One possible reason for the high retention rate amongst teachers was that there was an ongoing relationship with the teacher and researcher during the study as the Datalogging was collected from the dynamic soundfield approximately every month. Although the class teachers were linked to the classroom before the allocation of the intervention and control status, during the study they were not blind to the outcomes. This was a further limitation of this study. This means that drawing any conclusions about the effectiveness of the dynamic soundfield system from the teacher's questionnaire have to be treated with caution.

The AfE (InCAS) standardised test results completed at the start and end of the study were blind marked by the University of Durham's CEM and this is another strength of this research. As revealed in Chapter 2 one of the main shortcomings identified in the literature review was that non-standardised methods of assessments were often used, and the results were susceptible to response and expectation bias as they were graded by teachers or researchers not blind to the intervention. Another strength of the current study was the ability to analyse the PIPS data at the start and end of primary 1 to identify any variance between the groups in terms of attainment. For example, there were no significant differences at the end of primary 1 in early reading and phonological processing. There were differences at the end of the first year in school for early mathematics.

This study differs from previous research when considering whether dynamic soundfield would be beneficial to learners from the most and least deprived quintiles. One of the aims of the research design was to attract a significant sample of participants from SIMD 1 and 5. No previous research into dynamic soundfield had targeted this cohort and this was a strength of the study design. Based on the findings presented in Chapters 5, 6 and 7 it would also appear sensible to examine the merits of the dynamic soundfield system on learners from SIMD 2. Another limitation of the research is that the participants' classification of social deprivation status was on quintiles rather than deciles. This was done to ensure that there was a manageable sample size for the researcher. As the findings from the current study indicate there is a beneficial effect for adaptive amplification amongst learners from SIMD 1 quintile the aim of future research should be to investigate the intervention for learners from SIMD 1 decile.

One of the limitations of the current study was that none of the SIMD 1 learners attended a classroom that had excellent C_{50} properties, meaning that no conclusion can be drawn regarding the effects of dynamic soundfield on reducing the attainment gap in rooms with optimal early reflections. However, it can be hypothesised that learners with lower levels of attainment, that demonstrated significant improvements in good C_{50} classrooms would gain an additional advantage of a higher SNR provided by an excellent C_{50} environment. Future research should investigate the combined effects of excellent C_{50} classrooms and dynamic soundfield on closing the poverty-associated attainment gap between the most and least disadvantaged learners.

A final limitation of the current study to note is that the results on the effectiveness of dynamic soundfield relate to learners at the primary 3 stage of schooling. On the one hand, it would be reasonable to infer that younger learners may achieve a similar benefit of adaptive amplification as previous studies have shown that younger learners go through a period of maturation and development when listening in noise (Boothroyd, 1997). On the other hand, although it has been shown that young people in their teenage years are not as effective at listening in noise as adults (Johnson, 2000), it is unclear whether the later stages in primary would achieve a similar release from masking as the primary 3 learners. Therefore, it would appear sensible for future research to examine the effectiveness of dynamic soundfield on an older age group.

8.3 Contribution to the field

Young people from areas of social deprivation in the control classrooms were more affected by the noise created by learners moving around the school and the noise generated from other classrooms. Also, internal classroom noise produced by fidgeting and movement was identified as a significant distraction by SIMD 1 learners. This is a new finding, not revealed in previous studies. Previous research has established the relationship between classroom noise and the negative effects on attention, cognition, and attainment (Cohen et al., 1980, Shield and Dockrell, 2008, Klatte et al., 2013). Research has also shown that learners that are deaf, have English as an additional language or receive learning support in school are more affected by poor acoustics (Connolly et al., 2013, Bhang et al., 2018). The present study has extended existing knowledge by demonstrating that learners from areas of social deprivation are also more disproportionately affected than previously thought by this type of background and underlying noise. The findings from this study highlight the need for class teachers, and the leadership teams in schools to take into account the relationship between classroom noise and attainment when developing plans to tackle inequality in the Scottish education system.

The current study design only included enclosed school classrooms that offer the occupants greater acoustic privacy compared to an open plan environment. This leads to the other (rather obvious) finding that open-plan classrooms would be subject to greater levels of intrusive noise generated by the occupants in other teaching areas and movement around the school. A review of open plan schools found that noise from adjacent classes reduced speech intelligibility, distracted the attention of the learners and was a source of dissatisfaction for both learners and teachers (Shield et al., 2010). The results from this study emphasise the need for policymakers, planners, and practitioners to carefully consider managing distracting noise when planning and configuring school buildings. Future research should also investigate whether building open plan schools in areas of social deprivation is consistent with the Scottish Government's objective of closing the poverty-associated attainment gap.

The present study has extended existing knowledge by demonstrating the combined benefits of dynamic soundfield and early reflections (C_{50}) on educational attainment in areas of reading and developed ability. It is known that early reflections are important to improve speech intelligibility when the direct sound has been attenuated. Many situations in a contemporary classroom give rise to this, for example when the teacher is talking in the middle or back of the classroom, in group work where the talker is facing away from a section of the group or where the teacher is facing away from the class and towards an interactive whiteboard. In each of these scenarios, the learners are behind the person talking, which reduces access to direct sound and so the listener requires access to early reflections of speech sounds to enhance intelligibility (Sato et al., 2005, Bradley, 2009). The findings from this study are well aligned with past research which suggests that the primary acoustical design parameter for schools should not be reverberation times but maximising early reflections for speech. The results from this study concur with Bradley (2009) who recommends that the priority when designing a classroom should be to maximise the total energy in the direct sound and early reflections. The second priority would be to ensure that the long reverberation times are not excessive.

BB93 recommends that the T_{mf} for unoccupied classrooms for deaf learners should be 0.4s with no reference to a specified C_{50} value (Department of Education and Skills, 1993). These guidelines are supported by organisations supporting deaf learners such as the British Association of Teachers of the Deaf (BATOD) (BATOD, 2017). The absorption required to achieve short reverberation times may attenuate early reflections that improve the overall SNR in a classroom. Early reflections can provide an SNR of up to 9dB, whilst the doubling of RT_{60} would only increase long arriving sounds by 3dB (Bradley et al., 2003). The results from this study provide evidence for the combined effects of dynamic soundfield and good acoustics for speech clarity on educational outcomes. More significantly, they highlight the need for building standards for schools, including provision for the deaf to specify recommended levels of C_{50}. Using the rating scale from this research (Figure 26), excellent classrooms with C_{50} weighted values of 9dB would equate to an STI of 0.8. Combined with dynamic soundfield, this would contribute to an optimal SNR that would be appropriate for young learners.

The ISO 3382 provides objective data on the acoustical characteristics of a room (ISO 3382 1, 2009). The Aurora's Acoustical Parameter plug-in is a complete toolkit for acoustic measurements which complies with ISO 3382 and provides a large amount of data including C_{50}, strength (G), RT_{20}, RT_{30} and Early Decay Time (EDT) (Campanini and Farina, 2009). The plug-in can be used in conjunction with Audacity and so makes it an easy and cost-effective method to obtain a comprehensive acoustic fingerprint of a classroom. This plug-in, together with the rating-scale (Figure 26) was used to measure the speech clarity and reverberation times of the research classrooms and categorise the rooms from poor to excellent. The tests and rating scale in the current study can be used together as a tool to assess the suitability of teaching spaces for speech. By providing both C_{50} and RT_{30} data in a single, easy to use test it will allow professionals to ensure that the classroom provides the correct balance between early and long reflection energy.

Recent Scottish Government policy on school buildings has focused on the links between pedagogy and the learning environment. The overall aim is to reimage the traditional classroom into flexible learning spaces to facilitate individual and collaborative working. Interestingly, no reference is made to the acoustic properties of these flexible spaces or the need to establish acoustic privacy appropriate to the needs of vulnerable learners. The present study adds to the growing evidence base that shows that young people find attending to speech in noise more challenging than adults (Johnson, 2000, Connolly et al., 2013). This thesis has shown that learners from the most deprived quintile are even more vulnerable to the effects of background noise than the general population. Rather than designing a school to meet the needs of a curriculum, classrooms and the inclusive technology inside the schools should be tailored to meet the needs of the learner. There should be a closer association between architects, construction companies, the clients and the Scottish Government's educational priorities, such as the National Improvement Framework. School buildings should be designed to meet the aims of closing the poverty-associated attainment gap which should consider the detrimental effects of noise on attainment.

Headteachers and policymakers should become more familiar with the acoustic properties of their classrooms and the effect that classroom acoustics have on attainment. The present study developed a user-friendly rating scale for predicting the level of speech clarity and intelligibility in a classroom, based on speech-weighted C_{50} values. The test was able to differentiate between classrooms that were excellent, good, fair and poor for speech clarity. This rating scale can be used more widely in education as a measurement tool to establish the suitability of the classroom for speech clarity. When combined with the knowledge that noise has a detrimental effect on attention, speech intelligibility, and cognition, this may prove useful by allowing headteachers to locate young and vulnerable learners in rooms more appropriate to their needs. Furthermore, as the rating scale is referenced against the Speech Transmission Index, it may also provide audiologists useful data when fitting hearing aids to young deaf learners. Thus, it can be concluded that this rating scale can be used as a tool to assess classrooms for speech clarity for educational, health and research purposes.

To improve educational outcomes and reduce the poverty-associated attainment gap, policymakers should consider the installation of dynamic soundfield in early years classrooms. As discussed earlier in this thesis, external validity was established as the participants in this study are representative of both learners at the same primary 3 stage as the research cohort and all primary learners across Scotland. The current study extends existing knowledge and overcomes the substantial limitations of previous studies by demonstrating that overall attainment levels in developed ability, mental arithmetic, and reading were improved in standardised testing, blind marked by a third party. Furthermore, the results from the listening in noise questionnaire demonstrated the importance of dynamic soundfield in allowing young learners access to speech, even when noise levels were low. Dynamic soundfield also provided an advantage to hearing speech when visual cues were removed, and the direct sound was attenuated. Part of the Scottish Government's raising attainment policy should incorporate a school design-build that includes appropriate acoustics for speech clarity and technology that allows the adaptive amplification of the teachers' voice. This would need to be supported by appropriate training and maintenance of the dynamic soundfield system as the results obtained in this research were in schools where the systems were used more than of two-thirds of the school day.

The other finding was that SIMD 1 learners gained a significant benefit from dynamic soundfield in the curricular areas of developed ability and word decoding. Only learners from the most deprived quintile demonstrated a significant benefit for dynamic soundfield amplification in all the developed ability subtests. At follow-up, both non-verbal ability and picture vocabulary scores were higher than at baseline for SIMD 1 learners exposed to dynamic soundfield. Similar outcomes were observed for the word decoding subtest, one of the core skills required to be a successful reader. Another significant outcome of this research was the marked reduction in the poverty associated attainment gap in the subtests of non-verbal ability and word recognition. These findings concur with previous studies that show that young people require to have an enhanced SNR to achieve adult-like performance in noise (Bradley and Sato, 2004). These findings pave the way for the introduction of dynamic soundfield to improve learning outcomes in education. The Scottish Government (2016b) has a digital technology strategy to enhance learning and teaching and they should consider including dynamic soundfield in this policy.

The present study also demonstrated that young people find it challenging to hear the teacher easily in noise. Even when noise levels were low, the young learners found it more challenging to hear when visual cues were removed. Furthermore, in a modern classroom, there are several situations where the class teacher is not facing all of the students, which attenuates the direct sound and so early reflections are necessary to make speech accessible. Practitioners should consider routinely collecting the views of young people on their experience of listening in noise in the classroom. The noise questionnaire used in this study can be directly applied in schools as a measurement tool.

The reasons behind the persistent poverty-associated attainment gap are complex. Sosu and Ellis (2014) recommend more data-driven and evidenced-based approaches to identify the interventions that are effective. The findings from the current research suggest that policymakers should prioritise classrooms that provide good/excellent C_{50} values and the use of dynamic soundfield technology as part of a strategy to improve educational outcomes for all young people, including learners from the most socially deprived areas.

References

ALLEN, J. C. 2011. Sample Size Calculation for Two Independent Groups: A Useful Rule of Thumb. *Proceedings of Singapore Healthcare,* 20**,** 138-140.

AMICO, K. R. 2009. Percent total attrition: A poor metric for study rigor in hosted intervention designs. *American Journal of Public Health,* 99**,** 1567-1575.

ANDERSON, K. L. 1998. Listening Inventory For Education. Tampa: Educational Audiology Association.

ANDERSON, K. L. & GOLDSTEIN, H. 2004. Speech perception benefits of FM and infrared devices to children with hearing aids in a typical classroom. *Language, Speech, and Hearing Services in Schools,* 35**,** 169.

ANDERSON, S., WHITE-SCHWOCH, T., PARBERY-CLARK, A. & KRAUS, N. 2013. A dynamic auditory-cognitive system supports speech-in-noise perception in older adults. *Hearing research,* 300**,** 18.

ARBOGAST, T. L., MASON, C. R. & KIDD JR, G. 2002. The effect of spatial separation on informational and energetic masking of speech. *The Journal of the Acoustical Society of America,* 112**,** 2086-2098.

ARNOLD, P. & CANNING, D. 1999. Does classroom amplification aid comprehension? *British Journal of Audiology,* 33**,** 171-178.

ATILGAN, H., TOWN, S. M., WOOD, K. C., JONES, G. P., MADDOX, R. K., LEE, A. K. C. & BIZLEY, J. K. 2018. Integration of Visual Information in Auditory Cortex Promotes Auditory Scene Analysis through Multisensory Binding. *Neuron,* 97**,** 640-655.e4.

AZIZ, N. A. A., ALI, Z., NOR, N. M., BAHARUM, A. & OMAR, M. 2016. Modeling multinomial logistic regression on characteristics of smokers after the smoke-free campaign in the area of Melaka. Advances In Industrial And Applied Mathematics: Proceedings of 23rd Malaysian National Symposium of Mathematical Sciences (Sksm23), Johor Bahru, Malaysia (24–26 November 2015):AIP Conference Proceedings, 21 June 2016, Vol.1750(1).

BADDELEY, A. 2003. Working memory and language: An overview. *Journal of communication disorders,* 36**,** 189-208.

BAMFORD, J., UUS, K. & DAVIS, A. 2005. Screening for hearing loss in childhood: issues, evidence and current approaches in the UK. *Journal of medical screening,* 12**,** 119-124.

BATOD 2017. Acoustics and Technology. *BATOD Magazine*.

BEAR, M. F. 1996. *Neuroscience : exploring the brain,* Baltimore, Md. ; London, Baltimore, Md. ; London : Williams & Wilkins.

BEAR, M. F., CONNORS, B. W. & PARADISO, M. A. 2007. *Neuroscience*, Lippincott Williams & Wilkins.

BERG, F. S., BLAIR, J. C. & BENSON, P. V. 1996. Classroom acoustics: The problem, impact, and solution. *Language, Speech, and Hearing Services in Schools,* 27**,** 16.

BEST, V., GALLUN, F. J., CARLILE, S. & SHINN-CUNNINGHAM, B. G. 2007. Binaural interference and auditory grouping. *The Journal of the Acoustical Society of America,* 121**,** 1070-1076.

BEWICK, V., CHEEK, L. & BALL, J. 2004. Statistics review 8: Qualitative data - tests of association. *Critical care (London, England),* 8**,** 46.

BHANG, S.-Y., YOON, J., SUNG, J., YOO, C., SIM, C., LEE, C., LEE, J. & LEE, J. 2018. Comparing Attention and Cognitive Function in School Children across Noise Conditions: A Quasi-Experimental Study. *Psychiatry Investigation,* 15**,** 620-627.

BLANDY, S. E. & LUTMAN, M. E. 2005. Hearing threshold levels and speech recognition in noise in 7-year-olds: Niveles de umbrales auditivos y reconocimiento del lenguaje en ruido en niños de 7 años. *International journal of audiology,* 44**,** 435-443.

BOOTHROYD, A. 1997. Auditory development of the hearing child. *Scandinavian Audiology-Supplement Only,* 46**,** 9-16.

BOOTHROYD, A. Room acoustics and speech perception. Seminars in Hearing, 2004. Thieme Medical Publishers Inc, 155-166.

BRADLEY, J. A new look at acoustical criteria for classrooms. Proc. Inter-Noise, 2009.

BRADLEY, J., REICH, R. & NORCROSS, S. 1999. A just noticeable difference in C50 for speech. *Applied Acoustics,* 58**,** 99-108.

BRADLEY, J., SATO, H. & PICARD, M. 2003. On the importance of early reflections for speech in rooms. *Journal of the Acoustical Society of America,* 113**,** 3233-3244.

BRADLEY, J. S. & SATO, H. Speech intelligibility test results for grades 1, 3 and 6 children in real classrooms. Proceedings of the 18th Inter. Congress on Acoustics, ICA-2004, Kyoto, 2004.

BRADSHAW, P. 2011. *Growing Up in Scotland: Changes in child cognitive ability in the pre-school years*, Scottish Government Edinburg.

BRADSHAW, P., HALL, J., HILL, T., MABELIS, J. & PHILO, D. 2012. Growing up in Scotland: early experiences of primary school.

BREGMAN, A. S. 1993. Auditory scene analysis: Hearing in complex environments. *Thinking in sound: The cognitive psychology of human audition*, 10-36.

BRINK, D. 2010. *Essentials of statistics*, Bookboon.

BRONKHORST, A. W. 2000. The Cocktail Party Phenomenon: A Review of Research on Speech Intelligibility in Multiple-Talker Conditions. *Acta Acustica united with Acustica,* 86, 117-128.

BRUNGART, D. S. 2001. Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America,* 109, 1101-1109.

BRUNGART, D. S. & SIMPSON, B. D. 2002. The effects of spatial separation in distance on the informational and energetic masking of a nearby speech signal. *The Journal of the Acoustical Society of America,* 112, 664-676.

BRUNGART, D. S. & SIMPSON, B. D. 2004. Within-ear and across-ear interference in a dichotic cocktail party listening task: effects of masker uncertainty. *The Journal of the Acoustical Society of America,* 115, 301.

BRYMAN, A. 2015. *Social research methods*, Oxford : Oxford University Press.

BSA PROFESSIONAL PRACTICE COMMITTEE 2015. Recommended Procedure Pure-tone Air-conduction and Bone-conduction Threshold Audiometry With and Without Masking (2015). *The British Society of Audiology: Reading*.

BUSS, E., WHITTLE, L. N., GROSE, J. H. & HALL, J. W. 2009. Masking release for words in amplitude-modulated noise as a function of modulation rate and task. *The Journal of the Acoustical Society of America,* 126, 269.

CAMPANINI, S. & FARINA, A. 2009. *A new audicity feature: room objective acoustical parameters calculation module*, na.

CAMPBELL, D. T. & STANLEY, J. C. 1966. Experimental and quasi-experimental designs for research. *Handbook of research on teaching (NL Gage, Ed.)*, 171-246.

CARSON, M., GIESBRECHT, H. & PERRY, T. 2009. SURROUND SOUND IMPULSE RESPONSE. Measurement with the Exponential Sine Sweep; Application in Convolution Reverb. *Projet de fin d'études en Ingénierie électrique et informatique, Victoria (Texas, Etats-Unis), Université de Victoria*.

CENTRE FOR EVALUATING AND MONITORING 2012. Database Algorithims.

CENTRE FOR EVALUATING AND MONITORING 2014. Assessment for Excellence (AfE) P1-P7.

CHARLOTTE, C. & KATARINA, P. 2018. WHO Environmental Noise Guidelines for the European Region: A Systematic Review on Environmental Noise and Cognition. *International Journal of Environmental Research and Public Health,* 15**,** 285.

CHEN, H., COHEN, P. & CHEN, S. 2010. How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies. *Communications in Statistics - Simulation and Computation,* 39**,** 860-864.

CHERRY, E. C. 1953. Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America,* 25**,** 975-979.

CHOI, S., LOTTO, A., LEWIS, D., HOOVER, B. & STELMACHOWICZ, P. 2008. Attentional modulation of word recognition by children in a dual-task paradigm. *Journal of Speech, Language and Hearing Research,* 51**,** 1042.

COHEN, J. 1988. *Statistical power analysis for the behavioral sciences,* Hillsdale, N.J. ; Hove, Hillsdale, N.J. ; Hove : Lawrence Erlbaum.

COHEN, L. 2018. *Research methods in education,* New York, New York : Routledge.

COHEN, S., EVANS, G. W., KRANTZ, D. S. & STOKOLS, D. 1980. Physiological, motivational, and cognitive effects of aircraft noise on children: Moving from the laboratory to the field. *American Psychologist,* 35**,** 231-243.

COLEMAN, P. D. 1961. Experiments in hearing: Georg von Békésy. Translated and edited by E. G. Wever. New York, 1960, McGraw-Hill Book Company, Inc. Pp. 746, indexed. Price $25.00 (Book Review).

CONNOLLY, D., DOCKRELL, J., SHIELD, B., CONETTA, R. & COX, T. 2013. Adolescents' perceptions of their school's acoustic environment: The development of an evidence based questionnaire.(Article)(Report)(Survey). *Noise and Health,* 15**,** 269.

CONSOLIDATED STANDARDS OF REPORTING TRIALS. 2010. *The CONSORT Statement* [Online]. Available: http://www.consort-statement.org [Accessed 18/7/17].

COOKE, M. 2006. A glimpsing model of speech perception in noise. *The Journal of the Acoustical Society of America,* 119**,** 1562-1573.

COOKE, M., LECUMBERRI, M. G. & BARKER, J. 2008. The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *The Journal of the Acoustical Society of America,* 123**,** 414-427.

CRAGG, L., RICHARDSON, S., HUBBER, P. J., KEEBLE, S. & GILMORE, C. 2017. When is working memory important for arithmetic? The impact of strategy and age.(Research Article). *PLoS ONE,* 12**,** e0188693.

CRANDELL, C. C. & SMALDINO, J. J. 1996. Speech perception in noise by children for whom English is a second language. *American Journal of Audiology,* 5**,** 47-51.

CRIBBIE, R. A. & JAMIESON, J. 2004. Decreases in posttest variance and the measurement of change.

DA CRUZ, A. D., SILVERIO, K. C. A., RIBEIRO, V. V. & JACOB, R. T. D. S. 2016. Dynamic soundfield system impact on the teacher's voice: case report/Impacto do sistema de campo livre dinamico na voz do professor: estudo de caso.(Case reports. Texto en ingles). *Revista CEFAC: Atualizacao Cientifica em Fonoaudiologia e Educacao,* 18**,** 1260.

DADA, S., HUGUET, A. & BORNMAN, J. 2013. The Iconicity of Picture Communication Symbols for Children with English Additional Language and Mild Intellectual Disability. *Augmentative and Alternative Communication, 2013, Vol.29(4), p.360-373,* 29**,** 360-373.

DANCE, S., BACKUS, B. & MORALES, L. 2018. A methodology to objectively assess the performance of sound field amplification systems demonstrated using 50 physical simulations of classroom conditions. *Noise and Health,* 20**,** 77-82.

DARAI, B. 2000. Using sound field FM systems to improve literacy scores. *ADVANCE for Speech-Language Pathologists & Audiologists,* 10**,** 5.

DARWIN, C. J. 1997. Auditory grouping. *Trends in cognitive sciences,* 1**,** 327.

DEPARTMENT OF EDUCATION AND SKILLS. 1993. Acoustic Design of Schools. Available:http://webarchive.nationalarchives.gov.uk/20130401151715/https:/www.education.gov.uk/publications/eOrderingDownload/BB93-Acoustic_Design.pdf [Accessed 18/2/2014].

DILLON, H. 2001. *Hearing aids*, Thieme.

DOCKRELL, J. E. & SHIELD, B. 2012. The Impact of Sound-Field Systems on Learning and Attention in Elementary School Classrooms. *Journal of Speech, Language and Hearing Research,* 55**,** 1163.

DOCKRELL, J. E. & SHIELD, B. M. 2006. Acoustical barriers in classrooms: the impact of noise on performance in the classroom. *British educational research journal,* 32**,** 509-525.

DUBNO, J. R., AHLSTROM, J. B. & HORWITZ, A. R. 2008. Binaural Advantage for Younger and Older Adults with Normal Hearing. *Journal of Speech, Language, and Hearing Research,* 51**,** 539-556.

DURLACH, N. 2006. Auditory masking: need for improved conceptual structure. *The Journal of the Acoustical Society of America,* 120**,** 1787.

EDMONDS, B. A. & CULLING, J. F. 2006. The spatial unmasking of speech: Evidence for better-ear listening. *The Journal of the Acoustical Society of America,* 120**,** 1539-1545.

EDWARDS, B. 2007. The future of hearing aid technology. *Trends in amplification,* 11**,** 31-46.

EISENBERG, L. S., SHANNON, R. V., MARTINEZ, A. S., WYGONSKI, J. & BOOTHROYD, A. 2000. Speech recognition with reduced spectral cues as a function of age. *The Journal of the Acoustical Society of America,* 107**,** 2704.

ELLIOTT, L. L. 1979. Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *The Journal of the Acoustical Society of America,* 66**,** 651.

FERGUSON, C. J. 2009. An Effect Size Primer: A Guide for Clinicians and Researchers. *Professional Psychology: Research and Practice,* 40**,** 532-538.

FIELD, A. P. 2013. *Discovering statistics using IBM SPSS statistics : and sex and drugs and rock 'n' roll,* Los Angeles, Los Angeles : Sage.

FIFE COUNCIL 2009. Building Fife's Future.

FISCHBACHER, C. M. 2017. *Identifying "deprived individuals": are there better alternatives to the Scottish Index of Multiple Deprivation (SIMD) for socioeconomic targeting in individually based programmes addressing health inequalities in Scotland?* [Online]. Available: https://www.scotpho.org.uk/media/1166/scotpho140109-simd-identifyingdeprivedindividuals.pdf [Accessed 1/7/17].

FISHER, K. 2007. Building Excellence: Exploring the implications of the curriculum for excellence for school buildings' Dec 2007. Part.

FLETCHER, H. 1940. Auditory patterns. *Reviews of modern physics,* 12**,** 47.

FREYMAN, R. L., BALAKRISHNAN, U. & HELFER, K. S. 2004. Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *The Journal of the Acoustical Society of America,* 115**,** 2246-2256.

FROLENKOV, G. I. 2006. Regulation of electromotility in the cochlear outer hair cell. Oxford, UK.

GADE, A. C. 2014. Acoustics in halls for speech and music. *Springer handbook of acoustics.* Springer.

GARADAT, S. N. & LITOVSKY, R. Y. 2007. Speech intelligibility in free field: Spatial unmasking in preschool children. *The Journal of the Acoustical Society of America,* 121**,** 1047-1055.

GELFAND, S. A. 1998. Optimizing the reliability of speech recognition scores. *Journal of Speech, Language, and Hearing Research,* 41, 1088-1102.

GELFAND, S. A. 2016. *Hearing: An introduction to psychological and physiological acoustics*, CRC Press.

GENERAL TEACHING COUNCIL OF SCOTLAND. 2017. *Standards for Registration* [Online]. Available: http://www.gtcs.org.uk/professional-standards/the-standards/standards-for-registration.aspx [Accessed 2/6/17].

GILLOOLY, A. & RIDDELL, S. 2019. Working paper 1 (statistics update): An overview of statistics on SEN in England and ASN in Scotland.

GOSSELIN, P. A. & GAGNÉ, J.-P. 2010. Use of a Dual-Task Paradigm to Measure Listening Effort Utilisation d'un paradigme de double tâche pour mesurer l'attention auditive. *Revue canadienne d'orthophonie et d'audiologie,* 34, 43-51.

GREEN, S. B. 1991. How Many Subjects Does It Take to Do a Regression Analysis? *Multivariate Behavioral Research,* 26, 499-510.

GUINAN, J. J., SALT, A. & CHEATHAM, M. A. 2012. Progress in cochlear physiology after Békésy. *Hearing Research,* 293, 12-20.

GUMMER, A. W. & PREYER, S. 1997. Cochlear amplification and its pathology: emphasis on the role of the tectorial membrane. *Ear, nose, & throat journal,* 76, 151.

HAAS, H. 1972. The influence of a single echo on the audibility of speech. *Journal of the Audio Engineering Society,* 20, 146-159.

HAAS, J. P. 2012. Sample size and power. *AJIC: American Journal of Infection Control,* 40, 766-767.

HAE-YOUNG, K. 2013. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics,* 38, 52-54.

HAE-YOUNG, K. 2015. Statistical notes for clinical researchers: effect size. *Restorative Dentistry & Endodontics,* 40, 328-331.

HAIDICH, A. B. 2010. Meta- analysis in medical research. *Hippokratia,* 14, 29.

HALL, I., W, J., BUSS, E. & GROSE, J. H. 2005. Informational masking release in children and adults. *The Journal of the Acoustical Society of America,* 118, 1605-1613.

HALL, W. J., GROSE, H. J., BUSS, B. E. & DEV, B. M. 2002. Spondee Recognition in a Two-Talker Masker and a Speech-Shaped Noise Masker in Adults and Children. *Ear and Hearing,* 23, 159-165.

HALLAM, S., KIRTON, A., PEFFERS, J., ROBERTSON, P. & STOBART, G. 2004. Evaluation of Project 1 of the Assessment is for Learning Development Programme: Support for Professional Practice in Formative Assessment-Final Report.

HÄLLGREN, M. 2005. Hearing and cognition in speech comprehension. Methods and applications.

HART, B. & RISLEY, T. R. 1995. *Meaningful differences in the everyday experience of young American children*, Paul H Brookes Publishing.

HARTAS, D. 2011. Families' social backgrounds matter: Socio-economic factors, home learning and young children's language, literacy and social outcomes. *British Educational Research Journal,* 37**,** 893-914.

HATTIE, J., BIGGS, J. & PURDIE, N. 1996. Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research,* 66**,** 99.

HAZAN, V. & BARRETT, S. 2000. The development of phonemic categorization in children aged 6–12. *Journal of phonetics,* 28**,** 377-396.

HEENEY, M. F. 2007. *Classroom sound field amplification, listening and learning*, University of Newcastle NSW,, Australia.

HEIMAN, G. 2013. *Basic statistics for the behavioral sciences*, Cengage Learning.

HELFER, K. S. & FREYMAN, R. L. 2005. The role of visual speech cues in reducing energetic and informational masking. *The Journal of the Acoustical Society of America,* 117**,** 842-849.

HELLBRÜCK, J., FASTL, H. & KELLER, B. Effects of meaning of sound on loudness judgements. 2014.

HENDERSON, J. 2010. Curriculum for Excellence. *Nutrition Bulletin,* 35**,** 40-44.

HESKETH, A. 1986. HEARING IMPAIRMENT, AUDITORY-PERCEPTION AND LANGUAGE DISABILITY - BAMFORD,J, SAUNDERS,E. *Journal of the British Association of Teachers of the Deaf,* 10**,** 80-81.

HIBINO, H. & KURACHI, Y. 2006. Molecular and physiological bases of the K+ circulation in the mammalian inner ear. *Physiology (Bethesda, Md.),* 21**,** 336.

HINDMAN, A. H., WASIK, B. A. & SNELL, E. K. 2016. Closing the 30 Million Word Gap: Next Steps in Designing Research to Inform Practice. *Child Development Perspectives,* 10**,** 134-139.

HODGSON, M. 1994. UBC-classroom acoustical survey. *Canadian Acoustics,* 22**,** 3.

HORVAT, M., JAMBROSIC, K. & DOMITROVIC, H. Methods of measuring the reverberation time. 3rd Congress of the Alps Adria Acoustics Association, Graz–Austria, 2007. 27-28.

HOWARD, C. S., MUNRO, K. J. & PLACK, C. J. 2010. Listening effort at signal-to-noise ratios that are typical of the school classroom. *International journal of audiology,* 49**,** 928-932.

HUMES, W. 2013. Curriculum for excellence and interdisciplinary learning. *Scottish Educational Review,* 45**,** 82-93.

INSTITUTE OF ACOUSTICS 2015. Acoustics of Schools:A Design Guide.

ISO 3382 1 2009. *BS EN ISO 3382-1:2009 - Acoustics. Measurement of room acoustic parameters. Performance spaces*.

ISRAEL, G. 1992. Determining sample size: University of Florida Cooperative Extension Service. *Institute of Food and Agriculture Sciences, EDIS*.

JAMIESON, J. 2004. Analysis of covariance (ANCOVA) with difference scores. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology,* 52**,** 277.

JERRIM, J. 2013. The Reading Gap: The socio-economic gap in children's reading skills: A cross-national comparison using PISA 2009.

JOHNSON, B. & CHRISTENSEN, L. 2000. *Educational research: Quantitative and qualitative approaches*, Allyn & Bacon.

JOHNSON, C. E. 2000. Children's phoneme identification in reverberation and noise. *Journal of Speech, Language and Hearing Research,* 43**,** 144.

JOHNSTONE, P. M. & LITOVSKY, R. Y. 2006. Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults. *The Journal of the Acoustical Society of America,* 120**,** 2177-2189.

JONES, P. R., MOORE, D. R. & AMITAY, S. 2015. Development of Auditory Selective Attention: Why Children Struggle to Hear in Noisy Environments. *Developmental Psychology,* 51**,** 353-369.

KAHNEMAN, D. 1973. *Attention and effort*, Citeseer.

KIDD JR, G., MASON, C. R., RICHARDS, V. M., GALLUN, F. J. & DURLACH, N. I. 2008. Informational masking. *Auditory perception of sound sources.* Springer.

KIESSLING, J., PICHORA-FULLER, M. K., GATEHOUSE, S., STEPHENS, D., ARLINGER, S., CHISOLM, T., DAVIS, A. C., ERBER, N. P., HICKSON, L., HOLMES, A., ROSENHALL, U. & VON WEDEL, H. 2003. Candidature for and delivery of audiological services: special needs of older people. *International journal of audiology,* 42 Suppl 2**,** 2S92.

KILLION, M. C. & DALLOS, P. 1979. Impedance matching by the combined effects of the outer and middle ear. *The Journal of the Acoustical Society of America,* 66**,** 599-602.

KLATTE, M., BERGSTRÖM, K. & LACHMANN, T. 2013. Does noise affect learning? A short review on noise effects on cognitive performance in children. *frontiers in Psychology,* 4**,** 578.

KLATTE, M., HELLBRÜCK, J., SEIDEL, J. & LEISTNER, P. 2010a. Effects of classroom acoustics on performance and well-being in elementary school children: A field study. *Environment and Behavior,* 42**,** 659-692.

KLATTE, M., LACHMANN, T. & MEIS, M. 2010b. Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting. *Noise and Health,* 12**,** 270-282.

KLATTE, M., WEGNER, M. & HELLBRÜCK, J. 'Noise in the school environment and cognitive performance in elementary school children. Part B-Cognitive psychological studies'. Forum Acusticum 2005, 2005.

KNECHT, H., NELSON, P., WHITELAW, G. & FETH, L. 2002. Background noise levels and reverberation times in unoccupied classrooms: Predictions and measurements. *American Journal of Audiology,* 11**,** 65-71.

KUCHIRKO, Y. 2019. On differences and deficits: A critique of the theoretical and methodological underpinnings of the word gap. *Journal of Early Childhood Literacy,* 19**,** 533-562.

LEECH, N. L. 2005. *SPSS for intermediate statistics use and interpretation,* Mahwah, N.J., Mahwah, N.J. : Lawrence Erlbaum.

LEIBOLD, L. J., BONINO, A. Y. & FLEENOR, L. The importance of establishing a time course for typical auditory development. A Sound Foundation through Early Amplification: Fourth International Pediatric Conference, 2007. 6-8.

LEIBOLD, L. J. & BUSS, E. 2013. Children's Identification of Consonants in a Speech-Shaped Noise or a Two-Talker Masker. *Journal of Speech, Language, and Hearing Research,* 56**,** 1144-1155.

LIBERMAN, A. M. 2005. How Much More Likely? The Implications of Odds Ratios for Probabilities. *American Journal of Evaluation,* 26**,** 253-266.

LJUNG, R., SÖRQVIST, P., KJELLBERG, A. & GREEN, A.-M. 2011. Poor Listening Conditions Impair Memory for Intelligible Lectures: Implications for Acoustic Classroom Standards. *Noise Notes,* 10**,** 39-46.

LONG, J. S. 2014. *Regression models for categorical dependent variables using Stata*, College Station, Texas : Stata Press, StataCorp LP.

LUCAS, J. W. 2003. Theory-testing, generalization, and the problem of external validity. *Sociological Theory,* 21**,** 236-253.

MALTBY, M. T. 2002. *Principles of hearing aid audiology*, Whurr.

MANLEY, G. A., NARINS, P. M. & FAY, R. R. 2012. Experiments in comparative hearing: Georg von Békésy and beyond. *Hearing Research,* 293**,** 44-50.

MARSHALL, L. G. 1995. Speech intelligibility prediction from calculated C50 values. *The Journal of the Acoustical Society of America,* 98**,** 2845.

MASSIE, R. & DILLON, H. 2006a. The impact of sound-field amplification in mainstream cross-cultural classrooms: Part 1 Educational outcomes. *Australian journal of education,* 50**,** 62-77.

MASSIE, R. & DILLON, H. 2006b. The Impact of Sound-Field Amplification in Mainstream Cross-Cultural Classrooms: Part 2 Teacher and Child Opinions. *Australian Journal of Education,* 50**,** 78-94.

MASSIE, R., THEODOROS, D., MCPHERSON, B. & SMALDINO, J. 2004. Sound-Field Amplification: Enhancing the Classroom Listening Environment for Aboriginal and Torres Strait Islander Children. *Australian Journal of Indigenous Education,* 33**,** 47-53.

MAXWELL, L. E. & EVANS, G. W. 2000. The effects of noise on pre-school children's pre-reading skills. *Journal of Environmental Psychology,* 20**,** 91-97.

MCHUGH, M. 2009. The odds ratio: calculation, usage, and interpretation. *Biochemia Medica***,** 120-126.

MCHUGH, M. L. 2013. The chi- square test of independence. *Biochemia medica,* 23**,** 143.

MCKENDRICK, J., BARCLAY C, CARR C, CLARK A, HOLLES J, PERRING E & L., S. 2011. *Our Rural Numbers Are Not Enough: An independent position statement and recommendations to improve the identification of poverty, income inequality and deprivation in rural Scotland Glasgow: Rural Poverty Indicators Action Learning Set* [Online]. Available: http://www.scotpho.org.uk/index.php?option=com_content&view=article&id=1393:rurality-key-references-and-evidence&catid=235:rurality&Itemid=1893 [Accessed 1/7/17].

MCSPORRAN, E., BUTTERWORTH, Y. & ROWSON, V. J. 1997. Sound field amplification and listening behaviour in the classroom. *British educational research journal,* 23**,** 81-96.

MEHRGARDT, S. & MELLERT, V. 1977. Transformation characteristics of the external human ear. *The Journal of the Acoustical Society of America,* 61**,** 1567-1576.

MELBY-LERVÅG, M., LYSTER, S.-A. H. & HULME, C. 2012. Phonological Skills and Their Role in Learning to Read: A Meta-Analytic Review. *Psychological Bulletin,* 138, 322-352.

MERRELL, C. & TYMMS, P. 2004. *Diagnosing and remediating literacy problems using INCAS software: Identifying reading and spelling difficulties and providing help* [Online]. Available: https://intranet.westlothian.gov.uk/CHttpHandler.ashx?id=7270&p=0 [Accessed 16/11/2017].

MIDI, H., SARKAR, S. K. & RANA, S. 2010. Collinearity diagnostics of binary logistic regression model. *Journal of Interdisciplinary Mathematics,* 13, 253-267.

MIKULSKI, W. & RADOSZ, J. 2011. Acoustics of Classrooms in Primary Schools - Results of the Reverberation Time and the Speech Transmission Index Assessments in Selected Buildings. *Archives of Acoustics,* 36.

MILLETT, P. & PURCELL, N. 2010. Effect of Sound Field Amplification on Grade 1 Reading Outcomes Effet de l'amplification en champ libre sur les performances de lecture des élèves de première année. *Revue canadienne d'orthophonie et d'audiologie-Vol,* 34.

MOODLEY, A. 1989. Acoustic conditions in mainstream classrooms. *Journal of the British Association of Teachers of the Deaf,* 13, 48-54.

MOORE, B. C. 2003. *An introduction to the psychology of hearing*, Academic press San Diego.

MOORE, B. C. 2008. Basic auditory processes involved in the analysis of speech sounds. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 363, 947-963.

MOORE, B. C. J. 2007. *Cochlear hearing loss : physiological, psychological and technical issues,* Chichester ; Hoboken, N.J., Chichester ; Hoboken, N.J. : John Wiley & Sons.

NAKAGAWA, S. & CUTHILL, I. C. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews,* 82, 591-605.

NDCS 2001. Deaf Friendly Schools.

NELSON, P., KOHNERT, K., SABUR, S. & SHAW, D. 2005. Classroom noise and children learning through a second language: Double jeopardy? *Language, Speech, and Hearing Services in Schools,* 36, 219-229.

NEMES, S., JONASSON, J. M., GENELL, A. & STEINECK, G. 2009. Bias in odds ratios by logistic regression modelling and sample size. *BMC medical research methodology,* 9, 56.

NICE. 2014. Surgical management of otitis media with effusion in children. Available: http://www.nice.org.uk/nicemedia/live/11928/48420/48420.pdf [Accessed 15/11/2017].

NITTROUER, S. 2002. Learning to perceive speech: How fricative perception changes, and how it stays the same. *The Journal of the Acoustical Society of America,* 112**,** 711-719.

NITTROUER, S. & BOOTHROYD, A. 1990. Context effects in phoneme and word recognition by young children and older adults. *The Journal of the Acoustical Society of America,* 87**,** 2705.

NITTROUER, S. & BURTON, L. T. 2005. The role of early language experience in the development of speech perception and phonological processing abilities: Evidence from 5-year-olds with histories of otitis media with effusion and low socioeconomic status. *Journal of communication disorders,* 38**,** 29-63.

NOORDZIJ, M., TRIPEPI, G., DEKKER, F. W., ZOCCALI, C., TANCK, M. W. & JAGER, K. J. 2010. Sample size calculations: basic principles and common pitfalls. *Nephrology Dialysis Transplantation,* 25**,** 1388-1393.

O'NEILL, R., ARENDT, J. & MARSCHARK, M. 2014. Report from the Achievement and Opportunities for Deaf Students in the United Kingdom: from Research to Practice project. *University of Edinburgh. h^ p://bit. ly/1HsLEjG*

PATERSON, L., BLACKBURN, L. & WEEDON, E. 2019. The use of the Scottish Index of Multiple Deprivation as an indicator to evaluate the impact of policy on widening. *Paterson , L , Blackburn , L & Weedon , E 2019 , ' The use of the Scottish Index of Multiple Deprivation as an indicator to evaluate the impact of policy on widening ' , Scottish Affairs , vol. 28 , no. 4 , pp.414-433 . https://doi.org/10.3366/scot.2019.0296*.

PHONAK. 2014. *Dynamic SoundField for Education - Product Details* [Online]. Available:http://www.phonak.com/com/b2c/en/products/more_products/soundfield/product-details.html [Accessed 15/11/2016].

PICHORA-FULLER, M., PALMER, C. & SEEWALD, R. 2007. Audition and cognition: What audiologists need to know about listening. *Hearing care for adults***,** 71-85.

PLACK, C. 2012. Hearing pitch-right place, wrong time? *Psychologist,* 25**,** 892-894.

PLOMP, R. 1978. Auditory handicap of hearing impairment and the limited benefit of hearing aids. *The Journal of the Acoustical Society of America,* 63**,** 533.

POLIT, D. F. & BECK, C. T. 2010. Generalization in quantitative and qualitative research: Myths and strategies. *International journal of nursing studies,* 47**,** 1451-1458.

PRAJAPATI, B., DUNNE, M. & ARMSTRONG, R. 2010. Sample size estimation and statistical power analyses. *Optometry today,* 16**,** 10-18.

QUARTIERI, J., D'AMBROSIO, S., GUARNACCIA, C. & IANNONE, G. 2009. Experiments in room acoustics: modelling of a church sound field and reverberation time measurements. *WSEAS Transactions on Signal Processing,* 5**,** 126-135.

RANGANATHAN, P., PRAMESH, C. & AGGARWAL, R. 2017. Common pitfalls in statistical analysis: Logistic regression.(Statistics)(Report). *Perspectives in Clinical Research,* 8**,** 148.

REA, L. M. 2014. *Designing and conducting survey research : a comprehensive guide*, San Francisco, CA : Jossey-Bass.

REED, P. & WU, Y. 2013. Logistic regression for risk factor modelling in stuttering research. *Journal of fluency disorders,* 38**,** 88-101.

RICHARDSON, J. T. E. 2011. Eta Squared and Partial Eta Squared as Measures of Effect Size in Educational Research. *Educational Research Review,* 6**,** 135-147.

RÖNNBERG, J., RUDNER, M., FOO, C. & LUNNER, T. 2008. Cognition counts: A working memory system for ease of language understanding (ELU). *International journal of audiology,* 47**,** S99-S105.

RÖNNBERG, N., RUDNER, M., LUNNER, T. & STENFELT, S. 2014. Assessing listening effort by measuring short-term memory storage and processing of speech in noise. *Speech, Language and Hearing,* 17**,** 123-132.

ROSENBERG, G. G., BLAKE-RAHTER, P., HEAVNER, J., ALLEN, L., REDMOND, B. M., PHILLIPS, J. & STIGERS, K. 1999. Improving Classroom Acoustics (ICA): A Three-Year FM Sound Field Classroom Amplification Study. *Journal of Educational Audiology,* 7**,** 8-28.

ROY, K. P. & BROWNE, S. D. 2010. Classroom Acoustics and Green Schools.

SAINANI, K. L. 2011. Understanding odds ratios. *PM & R : the journal of injury, function, and rehabilitation,* 3**,** 263.

SATO, H. & BRADLEY, J. S. 2008. Evaluation of acoustical conditions for speech communication in working elementary school classrooms. *The Journal of the Acoustical Society of America,* 123**,** 2064.

SATO, H., BRADLEY, J. S. & MORIMOTO, M. 2005. Using listening difficulty ratings of conditions for speech communication in rooms. *The Journal of the Acoustical Society of America,* 117**,** 1157-1167.

SCHAARS, M., SEGERS, E. & VERHOEVEN, L. 2017. Word decoding development in incremental phonics instruction in a transparent orthography. *Reading and Writing,* 30**,** 1529-1550.

SCHNEIDER, B. A., DANEMAN, M. & MURPHY, D. R. 2005. Speech Comprehension Difficulties in Older Adults: Cognitive Slowing or Age-Related Changes in Hearing? *Psychology and Aging,* 20**,** 261-271.

SCHNEIDER, B. A., LI, L. & DANEMAN, M. 2007. How competing speech interferes with speech comprehension in everyday listening situations. *Journal of the American Academy of Audiology,* 18**,** 559-572.

SCHULZ, K. F. & GRIMES, D. A. 2002. Sample size slippages in randomised trials: exclusions and the lost and wayward. *The Lancet,* 359**,** 781-785.

SCOTCEN. 2015. *Growing up in Scotland Additional Support Needs of Children Aged 8* [Online]. Available: http://www.scotcen.org.uk/news-media/press-releases/2015/june/study-shows-boys-aged-eight-twice-as-likely-as-girls-to-need-additional-support-for-learning/ [Accessed 15/11/2017].

SCOTTISH GOVERNMENT. 2009. *The Education (Additional Support for Learning) (Scotland) Acts 2004 & 2009* [Online]. Available: http://www.gov.scot/Publications/2009/11/03140104/3 [Accessed 14/11/2016].

SCOTTISH GOVERNMENT. 2012a. *Estimated population of settlements by broad age groups* [Online]. Available: https://www.nrscotland.gov.uk/files//statistics/population-estimates/special-area/mid-2012-settlements/2012-pop-est-sett-local-main-tab2a.pdf [Accessed 23/7/17].

SCOTTISH GOVERNMENT. 2012b. Historical Pupil Census. Available: http://www.gov.scot/Topics/Statistics/Browse/School-Education/HistoricDatasets/Rolls-2006 [Accessed 15/8/2017].

SCOTTISH GOVERNMENT. 2014a. *Fife Council Area - Demographic Fact Sheet* [Online]. Scottish Government Available: https://www.nrscotland.gov.uk/files/statistics/council-area-data-sheets/fife-factsheet.pdf [Accessed 18/09/2018].

SCOTTISH GOVERNMENT. 2014b. *Scottish Index of Multiple Deprivation* [Online]. Available: http://22fa0f74501b902c9f11-8b3fbddfa1e1fab453a8e75cb14f3396.r26.cf3.rackcdn.com/simd_448749_v7_20121217.pdf [Accessed 16/07/2016].

SCOTTISH GOVERNMENT. 2015a. *Pupil Census* [Online]. Available: http://www.gov.scot/Topics/Statistics/Browse/School-Education/PubPupilCensus [Accessed 17/09/2016].

SCOTTISH GOVERNMENT. 2015b. *The Scottish Survey of Literacy and Numeracy* [Online]. Available: http://www.gov.scot/Topics/Statistics/Browse/School-Education/SSLN [Accessed 17/09/2017].

SCOTTISH GOVERNMENT. 2015c. *Teacher Census* [Online]. Available: http://www.gov.scot/Topics/Statistics/Browse/School-Education/teachcenssuppdata [Accessed 1/06/17].

SCOTTISH GOVERNMENT 2016a. Children's development at the start of school in Scotland and the progress made during their first school year: An analysis of PIPS baseline and follow-up assessment data. APS Group.

SCOTTISH GOVERNMENT 2016b. Enhancing Learning and Teaching Through the Use of Digital Technology. APS Group.

SCOTTISH GOVERNMENT. 2016c. *The Scottish Survey of Literacy and Numeracy* [Online].Available:http://www.gov.scot/Topics/Statistics/Browse/School-Education/SSLN [Accessed 18/08/2018].

SCOTTISH GOVERNMENT. 2018a. *National Improvement Framework* [Online]. Available: http://www.gov.scot/Topics/Education/Schools/NationalImprovementFramework [Accessed 02/01/2019].

SCOTTISH GOVERNMENT. 2018b. *Schools for the future:funding allocations* [Online]. Scottish Government, St. Andrew's House, Regent Road, Edinburgh EH1 3DG Tel:0131 556 8400 ceu@scotland.gsi.gov.uk. Available: http://www.scotland.gov.uk/Topics/Education/Schools/Buildings/funding [Accessed 05/03/2019].

SCOTTISH GOVERNMENT. 2018c. *Scottish National Standardised Assessments* [Online]. Available: https://standardisedassessment.gov.scot [Accessed 02/01/2019].

SEEMIS. 2017. *Education Management Information System* [Online]. Available: https://www.seemis.gov.scot/site3/ [Accessed 31/5/17 2017].

SHANNON, R. V., ZENG, F.-G., KAMATH, V., WYGONSKI, J. & EKELID, M. 1995. Speech recognition with primarily temporal cues. *Science,* 270**,** 303-304.

SHAW, E. A. G. 1974. The External Ear. *In:* KEIDEL, W. D. & NEFF, W. D. (eds.) *Auditory System: Anatomy Physiology (Ear).* Berlin, Heidelberg: Springer Berlin Heidelberg.

SHIELD, B., CONETTA, R., DOCKRELL, J., CONNOLLY, D., COX, T. & MYDLARZ, C. 2015. A survey of acoustic conditions and noise levels in secondary school classrooms in England. *The Journal of the Acoustical Society of America,* 137**,** 177.

SHIELD, B. & DOCKRELL, J. E. 2004. External and internal noise surveys of London primary schools. *The Journal of the Acoustical Society of America,* 115**,** 730.

SHIELD, B., GREENLAND, E. & DOCKRELL, J. 2010. Noise in open plan classrooms in primary schools: A review. *Noise and Health,* 12**,** 225-234.

SHIELD, B. M. & DOCKRELL, J. E. 2008. The effects of environmental and classroom noise on the academic attainments of primary school children. *The Journal of the Acoustical Society of America,* 123**,** 133.

SHINN-CUNNINGHAM, B. G. 2008. Object-based auditory and visual attention. *Trends in Cognitive Sciences,* 12**,** 182-186.

SHINN-CUNNINGHAM, B. G. & WANG, D. 2008. Influences of auditory object formation on phonemic restoration. *The Journal of the Acoustical Society of America,* 123**,** 295-301.

SLACK, M. K. & DRAUGALIS, J. R. 2001. Establishing the internal and external validity of experimental studies. *American Journal of Health System Pharmacy,* 58**,** 2173-2184.

SOSU, E. & ELLIS, S. 2014. Closing the attainment gap in Scottish Education.

SRINIVASAN, N. K., JAKIEN, K. M. & GALLUN, F. J. 2016. Release from masking for small spatial separations: Effects of age and hearing loss a). *The Journal of the Acoustical Society of America,* 140**,** EL73-EL78.

STEPHENSON, M. 2007. *The effect of classroom sound field amplification and the effectiveness of otoacoustic emission hearing screening in school-age children*, New Zealand Health Technology Assessment, Department of Public Health and General Practic, Christchurch School of Medicine and Health Sciences.

STOLLMAN, M. H., VAN VELZEN, E. C., SIMKENS, H. M., SNIK, A. F. & VAN DEN BROEK, P. 2004. Development of auditory processing in 6-12-year-old children: a longitudinal study: Desarrollo del procesamiento auditivo en niños de 6-12 años: un estudio longitudinal. *International journal of audiology,* 43**,** 34-44.

SULLIVAN, G. M. & FEINN, R. 2012. Using Effect Size-or Why the P Value Is Not Enough. *Journal of graduate medical education,* 4**,** 279.

SURESH, K. & CHANDRASHEKARA, S. 2012. Sample size estimation and power analysis for clinical research studies. *Journal of Human Reproductive Sciences,* 5**,** 7-13.

SUSSMAN, E., WONG, R., HORVÁTH, J., WINKLER, I. & WANG, W. 2007. The development of the perceptual organization of sound by frequency separation in 5–11-year-old children. *Hearing Research,* 225**,** 117-127.

SWEETOW, R. W. & HENDERSON-SABES, J. 2004. The case for LACE: Listening and auditory communication enhancement training. *The Hearing Journal,* 57**,** 32-35.

TIESLER, G. & OBERDÖRSTER, M. 2008. Noise — A Stressor? Acoustic Ergonomics of Schools. *Building Acoustics,* 15**,** 249-261.

TORGERSON, C. J., TORGERSON, D. J. & DIRECTOR, Y. T. U. 2013. Randomised trials in education: An introductory handbook. *London: Education Endowment Foundation*.

TORGERSON, D. 2008. *Designing randomised trials in health, education and the social sciences: an introduction*, Springer.

TURTUREAN, C. 2015. Who's Afraid of the Effect Size? *Procedia Economics and Finance,* 20**,** 665-669.

TYMMS, P. & HANNA, R. 2008. Interactive, diagnostic formative assessment on a large scale. *Paper Presented at IAEA conference*.

TYMMS, P., MERRELL, C. & HENDERSON, B. 2000. Baseline assessment and progress during the first three years at school. *Educational Research and Evaluation,* 6**,** 105-129.

TYMMS, P., MERRELL, C. & JONES, P. 2004. Using baseline assessment data to make international comparisons. *British educational research journal,* 30**,** 673-689.

ULFENDAHL, M. & FLOCK, Å. 1998. Outer hair cells provide active tuning in the organ of Corti. *Physiology,* 13**,** 107-111.

VAILLANCOURT, D. V., LAROCHE, D. C., GIGUÈRE, D. C. & SOLI, D. S. 2008. Establishment of Age-Specific Normative Data for the Canadian French Version of the Hearing in Noise Test for Children. *Ear and Hearing,* 29**,** 453-466.

VALENTE, D. L., PLEVINSKY, H. M., FRANCO, J. M., HEINRICHS-GRAHAM, E. C. & LEWIS, D. E. 2012. Experimental investigation of the effects of the acoustical conditions in a simulated classroom on speech recognition and learning in children. *The Journal of the Acoustical Society of America,* 131**,** 232.

VANVOORHIS, C. W. & MORGAN, B. L. 2007. Understanding power and rules of thumb for determining sample sizes. *Tutorials in Quantitative Methods for Psychology,* 3**,** 43-50.

VICKERS, D. A., BACKUS, B. C., MACDONALD, N. K., ROSTAMZADEH, N. K., MASON, N. K., PANDYA, R., MARRIAGE, J. E. & MAHON, M. H. 2013. Using Personal Response Systems to Assess Speech Perception Within the Classroom: An Approach to Determine the Efficacy of Sound Field Amplification in Primary School Classrooms. *Ear and hearing*.

VISWANATHAN, N., KOKKINAKIS, K. & WILLIAMS, B. T. 2016. Spatially separating language masker from target results in spatial and linguistic masking release. *The Journal of the Acoustical Society of America,* 140**,** EL465-EL470.

WÅLINDER, R., GUNNARSSON, K., RUNESON, R. & SMEDJE, G. 2007. Physiological and psychological stress reactions in relation to classroom noise. *Scandinavian Journal of Work, Environment & Health,* 33**,** 260-266.

WATSON, C. S. 2005. Some Comments on Informational Masking. *Acta Acustica united with Acustica,* 91**,** 502-512.

WIGHTMAN, F., KISTLER, D. & BRUNGART, D. 2006. Informational masking of speech in children: auditory-visual integration. *The Journal of the Acoustical Society of America,* 119**,** 3940.

WIGHTMAN, F. L. & KISTLER, D. J. 2005. Informational masking of speech in children: Effects of ipsilateral and contralateral distracters. *The Journal of the Acoustical Society of America,* 118**,** 3164-3176.

WIGHTMAN, F. L., KISTLER, D. J. & O'BRYAN, A. 2010. Individual differences and age effects in a dichotic informational masking paradigm a). *The Journal of the Acoustical Society of America,* 128**,** 270-279.

WILSON, W. J., MARINAC, J., PITTY, K. & BURROWS, C. 2011a. The use of sound-field amplification devices in different types of classrooms. *Language, Speech, and Hearing Services in Schools,* 42**,** 395.

WILSON, W. J., MARINAC, J., PITTY, K. & BURROWS, C. 2011b. The use of sound-field amplification devices in different types of classrooms.(Report)(Report). *Language, Speech, & Hearing Services in Schools,* 42**,** 395.

WORLD HEALTH ORGANIZATION 2007. *International Classification of Functioning, Disability, and Health: Children & Youth Version: ICF-CY*, World Health Organization.

YAP, B. W. & SIM, C. H. 2011. Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation,* 81**,** 2141-2155.

YEE CHOI, C. & MCPHERSON, B. 2005. Noise levels in Hong Kong primary schools: Implications for classroom listening. *International Journal of Disability, Development and Education,* 52**,** 345-360.

YOUNG, E. 2007. Physiological Acoustics. *In:* ROSSING, T. D. (ed.) *Springer Handbook of Acoustics.* New York, NY: Springer New York.

ZHANG, J. & YU, K. F. 1998. What's the Relative Risk?: A Method of Correcting the Odds Ratio in Cohort Studies of Common Outcomes. *JAMA,* 280**,** 1690-1691.